American Psychological Assoc.
Committee on the Standardizing
of Procedure in Experimental
Tests

Association tests...

Stack copy

Association Ideas
Psychology Experimental
and Physiological

Vol. XIII
No. 5

PSYCHOLOGICAL REVIEW PUBLICATIONS

DECEMBER, 1911
Whole No. 57

THE

Psychological Monographs

Association Tests

Being a Part of the Report of the Committee of the American Psychological Association on the Standardizing of Procedure in Experimental Tests

By
R. S. WOODWORTH
Columbia University

and

FREDERIC LYMAN WELLS
MacLean Hospital

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, N. J.,
BALTIMORE, MD. and LANCASTER, PA.

Agents: G. E. STECHERT & CO., London (2 Star Yard, Carey St., W. C.); Leipzig (Hospital St., 10); Paris (76 rue de Rennes)

CONTENTS.

ASSOCIATION TESTS

I. SCOPE OF THE WORK

The present paper forms part of the report of a Committee of the American Psychological Association, appointed in 1906, "to act as a general control committee on the subject of measurements." The Committee consists of Professor James R. Angell, chairman, and Professors Judd, Pillsbury, Seashore and Woodworth. This Committee was authorized to organize sub-committees and to secure the assistance of other members of the Association. A sub-committee on association tests was appointed. consisting of the present authors, and the present paper, the report of this sub-committee, is to be regarded as a supplement to the "Report of the Committee of the American Psychological Association on the Standardizing of Procedure in Experimental Tests," published in 1910 as No. 53 of the *Psychological Monographs*.

The Association entrusted to its Committee two general lines of work: "first, the determination of a series of group and individual tests, with reference to practical application, and second, the determination of standard experiments of a more technical character." The sub-committee on association tests has confined itself to the first of these two lines of work. Leaving aside the more elaborate procedure, with chronoscope and lip key, we have fixed our attention on the "tests" so frequently employed in individual and pathological psychology for determining the speed and quality of association. Tests are needed which shall not require elaborate apparatus nor the expenditure of much time on the part of the individual tested. Many such tests are in use; these we have attempted to sift and, where possible, improve. The manner of giving the tests has varied from one experimenter to another; and we have attempted to ascertain the advantages and defects of the different procedures, and to make recommendations accordingly.

The efforts of a standardizing committee are likely to be regarded with disfavor and apprehension in many quarters, on the ground that the time is not yet ripe for stereotyping either the test material or the procedure. It may be felt that what is called for, in the present immature condition of individual psychology, is, the rather, free invention and the appearance of as many variants as possible. Let very many tests be tried, each new investigator introducing his own modification; and then, the worthless will gradually be eliminated and the fittest will survive. Admitting the general justice of this point of view, we still believe that work such as is here undertaken may be of service in two ways.

First, we hope that the tests herein recommended may find application where no special reason exists for the introduction of a new test. Often appeals for tests of proved value are heard from those who desire to study individual, race, sex, child or pathological psychology—from investigators who have not the time or inclination to devise new tests, and who, moreover, wish to be able to compare their results on one class of subjects with results already obtained on other classes. If every fresh student employs new tests, the incomparability of the results entails much wasted effort. Individual and class psychology is, almost of necessity, a cooperative enterprise. The advantages to be hoped from standardization are much the same here as in the field of anthropometry.

Second, it can scarcely fail to be true in psychology as in all other sciences that a full study of the *methods,* though too time-consuming and too remote from final results to be attractive at the start, is certain to lead to more reliable results in the long run. In the field of association—aside from the more technical experiments in memory—the methods have not been much subjected to the kind of experimental criticism which is here attempted. Usually the investigator has pressed forward to the solution of his problem, devising tests that seemed suitable to his purpose, and then abiding by them. Our concern being, on the contrary, exclusively with the tests themselves, we have sought for evidences of their relative value, relying at first on the ex-

perience of previous investigators, but in the last resort on renewed experiment with this end in view.

The tests which we have thus selected are in some degree analogous to "tested reagents" in chemistry. They make no claim, indeed, to be "chemically pure;" that is to say, they can not be guaranteed to give a true measure of every individual tested. Any mental test is sure to be vitiated in some cases either by peculiarities of an individual's training and information, or by the accidental variations to which mental performance is subject from moment to moment. These sources of error exist in all measurements of intellectual abilities. In the face of such difficulties, some investigators have felt it necessary to retreat from a quantitative attack on individual psychology; while others, more hopeful, have sought to neutralize the error of the single measurement by statistical methods. In the study of class differences, they have relied on averages from large groups; and in the study of correlations, they have endeavored to correct for the attenuation resulting from chance errors in the single measurements. But either reliance on the averages of large groups or reliance on Spearman's attenuation formulae is a reliance on probability, and therefore sure to be justified in the long run, but equally sure to be treacherous somewhere or other. Certainly, therefore, it is wise to eliminate from the tests all possible sources of error; though other sources of error still remain, yet for every defect eliminated there is an increase in the reliability of the individual measure, and so of the final result. Now most of the tests hitherto employed involve sources of error which can be eliminated once they are detected in practise. Many of these sources of error are little details in the construction of the tests; for example, one or two of the words selected as stimuli may have been ambiguous, or unfamiliar to many subjects. Our work has very largely consisted in attention to such details; and while we cannot hope to have attained perfection of detail, we are sure that we have taken some steps in that direction.

There is general agreement, in practise, as to what shall be included under the heading of association tests. There is the "free association" test, and the various tests of "controlled as-

sociation." In theory, indeed, it is difficult to draw any sharp line between association and memory, or intelligence, or reaction time. Every mental test involves association; but, in practise, the association test is regarded as limited to rather simple intellectual performances, and thus is distinguished from more complex tests of intelligence. On the other hand, the association between stimulus and response which comes into play in the simple reaction, or in the discriminative reaction, is simpler and quicker-acting than that of the "associative reaction." The line is, however, not sharply defined, and we have included in our list of tests one or two (as the "number-checking test") which might be more properly classed under the head of discrimination.

The distinction between association and memory experiments is even harder to draw in theory, though in practise the two are well enough distinguished. In the typical memory experiment new associations are first formed and later examined as to their strength; whereas the association experiment deals with association already formed, and does not control the process of their formation. Herein appears an obvious deficiency of the association experiment as compared with the standard experiments on memory. The memory experiment deals with a limited system of associations, formed specially for the purpose of the experiment and under controlled conditions. The association experiment dips into the general mass of the individual's associations, formed at various times and under varying conditions, with varying degrees of frequency, recency, vividness, emotional and intellectual value; and all these conditions vary from one individual to another. An experiment in the formation of entirely new associations gives all individuals an equal start; but a test dealing with previously formed associations can not hope to be perfectly fair. It aims, let us say, to give a measure of the speed of the individual's associative processes; but what it actually measures is, to a large extent, the familiarity of the particular associations called for, and the freedom of these associations from external interferences.

In the face of these difficulties, the association test may still prove of value. It may serve any one of at least three purposes,

and must be specifically adapted to the purpose which it is required to serve.

(1) A measure of the speed of formation of new associations. Such a test is indistinguishable from an experiment in memory or practise; but we have included one such, the "substitution test."

(2) Mental Diagnosis. Here the fact that the same association may have very different values in different individuals is fully recognized, and the object in view is to determine the value of a given association in the individual. Besides the emotional value, of which use is made in "psychoanalysis", the interest of a particular association may be the object of inquiry, as in *"Tatbestandsdiagnostik."* Also, the individual's familiarity with a certain sort of subject matter, or with a given form of logical relation, may be the thing measured. Thus the psychoanalytic viewpoint in association tests can be used, not only for the diagnosis of disturbing ideas and complexes, and for the detection of concealed knowledge, but also for showing the lines of thought with which an individual is conversant, and the sort of relationships along which his mind habitually moves. These uses of the association tests often require such close adaptation of the experimental material to the special object in view that they cannot easily be provided for by a standardized series of tests.

(3) A measure of mental alertness. The speed of an associative reaction depends not only on the strength of the associative tendency called into action—and thus on the previous training of that association—but also on the "determining tendency" or "adjustment" or "set of mind." In controlled association, the speed of the reaction depends on the efficiency of the control. In free association, also, a certain adjustment is required in order that the stimulus may call out a quick response: there must be a receptive attitude, a repression of any train of thought that would interfere with the speedy apprehension of the meaning of the stimulus; and there must also be an adjustment to give prompt expression to the first idea suggested by the stimulus. In a test of either free or controlled associations, calling for a series of responses in quick succession to a series of stimuli, the speed

of the performance depends on maintaining the proper adjustment throughout the series, in opposition to the many interfering tendencies generated by the successive stimuli. Periods of confusion are apt to occur in the course of such a series; and when they occur they impede the action of even well-trained associations. One cause of such periods of confusion, as has been abundantly pointed out by Jung and his school, lies in the emotional value of certain stimulus words; but that this is by no means the only cause of confusion is made evident in the color naming and similar tests, in which the same few stimuli are repeated many times in chance order. The associations required are here thoroughly familiar, and usually operate with great promptness; but at times they refuse to act properly, so that, in the midst of a series of rapid reactions, delayed and even false reactions occur to the same stimuli. The confusion here is sometimes due to wandering of the attention from the work in hand; but at other times it seems to be due to interferences generated by the performance itself. Whatever may be the cause of confusion in each particular case, efficiency in the test requires such a degree of control as will eliminate the confusion. Periods of confusion are but extreme manifestations of inefficient control; in a minor degree, the inefficiency of one individual in comparison with another is shown by uniform slowness of response.

In order, however, to make the association tests a measure of efficient mental control, it is necessary that the associations demanded shall be equally familiar to the individuals compared. In strictness, it is impossible to make sure of this; for the experimenter has no sufficient knowledge of the frequency, recency, etc., of the training which the several associations have received. The best that can be done is to call only for such associations as are familiar to all, or at least to the class of individuals to be tested.

Regarded as a measure of mental alertness or efficiency of control, the association test should be susceptible of standardization; and the efforts of the sub-committee have accordingly been mostly directed to this end. We have in every case but one—the Kent-Rosanoff experiment—sought for tests in which the

speed of association could properly be taken as the measure of efficiency—tests from which the question of the *quality* of the responses could be practically eliminated. To this end we have sought to determine, usually by experiment, what associations are so generally familiar as to be fair material for a test of individual differences in speed of association. We have also studied different methods of administering these tests, with a view to contributing towards uniformity of procedure; and we have, finally, endeavored to furnish average results obtained by these tests with one class of subjects, namely young adults of fair to good education.

During the progress of our work, several important contributions to the subject have appeared, of which two should be specially mentioned, those of Whipple[1] and of Whitley.[2] The aim and apparently also the method of Professor Whipple in preparing his lists of tests are the same as those of the present report; but the scope of his work is much more inclusive, and the present paper therefore represents a more intensive study of a limited field. The lists of tests here offered may be regarded as supplementing Whipple's list at a point where it is not especially full nor especially standardized.

Dr. Whitley's work is concerned very largely, though not exclusively, with association and similar tests; and her purpose is the same as ours, namely, to test the tests, and determine by experiment which are better and which worse. Her methods are however different from ours, in that, while we have been principally concerned with the details of each test, seeking to eliminate defects and sources of error, she has taken a large number of tests, as they stood, and compared the results obtained by their use. She has tried many similar tests on the same subjects, and has moreover repeated the same test a number of times, and then has evaluated the tests by the following criteria: (1) the better tests should not show rapid improvement with practise, for very rapid improvement indicates that some device for dealing with the test, or some adaptation to the conditions of the test, is of

[1] Manual of Mental and Physical Tests, Baltimore, 1910, pp. 254-270, 312-343.
[2] An Empirical Study of Certain Tests for Individual Differences. *Archives of Psychology*, No. 18, 1911.

prime importance; and as some subjects may chance to hit upon the adaptation or device at once, and others not, the first trial is likely to assign an individual a false position in the function designed to be measured; (2) the best test should show only a small variation in repeated trials (after the practise effect is allowed for), for the greater the variability, the less reliable is the single trial or the average of a few trials; (3) the best of a number of similar tests is that which correlates most closely with the average of them all, for this test represents the fairest sampling of the group of similar mental performances which it is desired to measure. In point of method, then, Dr. Whitley's work and ours are complementary; for a good test must both be free from minor defects, and must serve to indicate the efficiency of a function somewhat broader than that of dealing with the exact material used in the test. In regard to results, it is not easy to compare the two pieces of work, so much depends on the particular tests examined; but we find agreement at several points. Dr. Whitley, like ourselves, finds the use of written responses inadmissible in a test for speed of association; her results also tend to give the preference to the use of easy and simple material, such as we have adopted; and some of the tests which came out best in her comparison—such as an "easy opposites" test, a "first idea" test, a letter-checking and a form-checking test, and a form-naming test—are very similar to some included in our list.

II. QUESTIONS OF PROCEDURE

I. THE FORM OF RESPONSE.

Where the time of each single reaction is taken, as in the classical experiments on association time, the response has almost always been a spoken word, and the apparatus has measured the time to the beginning of the vocal utterance. But in tests which have measured the time, not of each single reaction, but of a continuous series of reactions, several forms of response have been used. Spoken words, written words, written letters, written Arabic numerals, and strokes of the pencil, checking or cancelling some of the (visual) stimuli, have all been used in different tests. In a test of the speed of any mental process, it is clear that the motor expression necessary for experimental purposes should require as little attention as possible and occupy as little time as possible. None of the above mentioned forms of response require much attention from an educated subject, but speech and cancellation have some advantage in this respect over writing. In respect to the time occupied by the movement, also, writing is at a disadvantage. The different times occupied by these various sorts of motor expression can be judged from the following results, obtained from two educated subjects:

Time for reading (either aloud or silently) a column of 20 disconnected letters or Arabic numerals..	6-7	sec.
Time for reading (either aloud or silently) a column of 20 short words, with a total of 22 syllables..	6-7½	sec.
Time for copying 20 one-place numbers.....	10-11	sec.
Time for copying 20 disconnected letters.....	12-13	sec.
Time for copying 20 short words, containing a total of 80 letters.....	27-35	sec.
Time for cancelling each of a list of 20 letters or words	6-8	sec.

The oral response, and the cancelling movement, have therefore a great advantage even over the writing of numerals.

So slow a process as the writing of words could never be thought of as a suitable form of response, were it not for the fact, that when a series of stimuli, such as a column of numbers to be added or a list of words to which synonyms are required, is presented together, the perceptive, associative and motor processes overlap; while the subject is writing the response to the first stimulus, he is already dealing with the second stimulus. If therefore the motor response is such as to occupy little time in comparison with the associative process, the overlapping brings it about that the time for the series of responses is nearly identical with the time of the associative processes involved; but if the motor response takes a much longer time than the associative process, the time of the series, because of overlapping, is nearly identical with the time of the motor processes. Overlapping causes a disappearance of motor time in the first case, and of the association time in the second case. So time-consuming a movement as writing can only be used as an idex of the speed of association when the associations themselves are much more difficult and slow than those which are customary in mental tests.

With all this admitted, written responses might still find a defender, on the ground that the writing should be delayed by any halt in the associative process, so that, on the average, the longer the time required to write the list of responses, the slower must be the association. This is probably true; but it does little to weaken the objection to written responses. For, first, if only one individual is considered, or only individuals having the same speed of writing,—and if, also, the various words to be written are suitably adjusted as to length—then the longer writing time indicates the slower association, indeed, but the indication is far from sensitive, and fails altogether below a certain limit. Thus, for example, the associations involved in reading a list of words, and those involved in naming colors, are both too rapid to be measured by aid of written responses. The results of one well-trained subject may be given. To react to a series of 20 patches of color by speaking the names required 12 seconds; to read the 20 printed color-names required but 6 seconds; but to write the names, either in response to the colors or

in response to a list of the names, required in each case 28 seconds. Here written responses conceal very considerable difference in speed of association. Again, in case of the "opposites" test, a subject reacted to a list of twenty very familiar stimuli, by speaking the opposites, in 15 seconds; to a slightly less familiar set in 22 seconds; to a list of the response words, by reading them, in 6.5 seconds; but to copy the words from the list required 29 seconds; to write the responses to the easier set required 31 seconds, and to the harder set 30 seconds. Thus written responses entirely conceal the differences in speed of associations, provided only the association time is not over one second; and that even without regard to variations in the speed of writing. When however different individuals are to be compared, the speed of writing must be considered; and as this speed varies at least in the ratio of 2 to 1, even in educated adults, and as moreover, there is no close correlation (as we have found) between the speed of writing and the speed of association among educated subjects, it is clear, in conclusion, that conditions can scarcely be so favorable as to justify the use of written words as responses in any test of individual differences in speed of association.

The case is not quite so unfavorable with the writing of single letters or one-place numbers. For example, it is easier to respond to a letter by giving the following letter than by giving the preceding letter; and this difference appears in either oral or written responses. (One subject, 2 trials, list of 20 letters: Preceding letter: oral, 32 sec.; written, 35 sec. Following letter: oral, 20 sec.; written, 25 sec.) The writing of single letters or numerals is an admissible form of response when the association time is over a second—provided the individuals tested are accustomed to rapid writing.

2. MEASUREMENT OF ASSOCIATION TIME.

As already remarked, the purpose for which the present set of tests is designed excludes the use of elaborate apparatus and therefore of the chronoscope and lip key. The custom of many

students of association time, in clinical and similar work, is to employ the stop watch, starting the watch together with the spoken stimulus word, and stopping it on hearing the beginning of the response. This procedure would seem to include the experimenter's simple reaction time (probably 150-200 σ under the conditions) in the measured time. Moreover there is no guarantee that the watch is started precisely together with the giving of the stimulus; a degree of error must be expected here; and moreover, the fifth of a second of the stop watch is scarcely to be called a fine unit. In spite of these objections, the use of the stop watch appears to be justified in practise, especially since the variation in association time is so great that significant differences can usually be established even with a rough procedure.

Another procedure, much employed when the chronoscope can not be used, is to expose (visually) a whole list of stimuli, and to require the subject to react to these in succession and without delay between the separate reactions. The time is then taken, not for the single reactions, but for the whole series. As the time necessary for reacting to the whole series is usually at least 10 seconds, and often much greater, the deficiencies of the stop watch are not serious in this procedure. As indicated in the preceding section, when the motor reaction requires little time or attention, the overlapping of motor and central processes brings it about that the time of such a series of responses is essentially central time. If therefore the total time of the series of reactions be divided by the number of responses in the series, the quotient should give the average association time.

It would seem possible, indeed, that overlapping should accomplish more than this, and make the average association time, computed as just described, considerably less than that obtained with single stimuli. Cattell[1] found that a series of disconnected words could be read at a rate of 200 σ per word, whereas the reading time for an isolated word was 360 σ. But in even slightly more difficult reactions, such as naming presented colors, this shortening of the reaction time, when a series of stimuli is presented together, does not appear; but the average time comes out at

[1] Wundt's *Philos. Studien*, 1885, 2, 635.

from 600 to 1200 σ. The same holds good for such associations as are involved in the opposites test. We have tested five individuals with the same stimulus words, first singly, and then, several months later, in lists. Though the first of these tests should have made the responses somewhat more familiar, only one of the five subjects reacted more quickly to the words in lists than separately; two subjects reacted more slowly to the words in lists, and two showed no marked or consistent difference. On the whole, the average time as obtained by timing lists of associative responses is no less, and probably somewhat greater, than that obtained from separate reactions.

Some explanation is demanded by the failure of overlapping to hasten the reaction to a series of stimuli. The explanation is probably found in interferences generated in the course of a rapid series of associations. Many associative tendencies are partially aroused by each stimulus word, and when no interval elapses between the successive reactions, the tendencies generated by the earlier members of the series must be held in check in order to give free play to the associations required by the later stimuli. Irrelevant associations enter and tend to impede the progress of the reactions. Introspection makes this view seem probable, for often the subject is conscious that trains of thought, started by the earlier stimuli, must be repressed in order to do justice to the later stimuli. Sometimes the response made to a stimulus is not wholly satisfactory to the subject; sometimes a second response to the same stimulus is suggested immediately after the first has been spoken; sometimes an interesting idea or disturbing emotion is suggested by a stimulus or by the response made to a stimulus. All such interferences die away with the lapse of a few seconds between the stimuli; but are present in full force when no interval is allowed. Success in dealing rapidly with a series of unrelated stimuli requires a higher degree of control than success in dealing with isolated stimuli.

This interpretation of the list or serial test is borne out by the following experiment.[1] The subject had before him a list of

[1] We are indebted to Mr. Franklin B. Pedrick for collaboration in this experiment. It is intended to present elsewhere a fuller report bearing on the question of fatigue within brief periods of mental work.

20 stimulus words, to which he reacted in quickest possible succession; but the experimenter, instead of timing simply the whole list, took the time for each reaction, or, at least, the interval between each two successive reactions. This was accomplished by bringing side by side on a revolving drum a Jacquet chronograph marking fifths of seconds and an electromagnetic marker connected with a telegraph key on which rested the experimenter's finger. The experimenter pressed the key on exposing the list to view, and then on hearing the beginning of each successive response of the subject. Thus a record of the distribution of time through the series of responses was obtained, having an accuracy somewhat superior to that obtained ordinarily with the stop watch. The method can not be employed where the series of responses is very rapid and regular (as in naming colors), for then a rhythmic tendency dominates the experimenter's hand; but when the intervals between responses vary irregularly from 0.4 to 2 or more seconds, the method is perfectly feasible. Nine subjects were so tested, each reacting to 20 lists of 20 words. The instructions called for supraordinate, subordinate concepts, etc., the task remaining the same through each list of 20 stimulus words. (The experiment was at the same time designed to indicate the comparative difficulty of the stimulus words, and so to aid in selection of the best lists.)

In combining the results obtained from several lists and from several individuals, with the object of determining the general distribution of time throughout the list, difficulty arises from the inequal difficulty of the lists and from the unequal speed of the individuals. If the times for all the first reactions are simply averaged, and so for all the second reactions, etc., the general tendency is obscured by the extraneous variations so introduced. We therefore proceeded as follows: Taking one individual's performance in response to one list, we determined the distribution of time throughout this one list, by first determining the average time of these 20 reactions, and the average deviation of the reactions, and then expressing the time of each reaction as + or − (according as the time of this reaction was greater or less than the average time of the twenty) such and such a per cent

of the average deviation of the reactions in that list. For example, a mark of −50 meant that the time of a reaction was 50 per cent of the average deviation less than the average time for the list. The same process was repeated with each of the 20 lists; and the marks so obtained were averaged for each position within a list. Thus an average of + 50 for an individual in the first place meant that his first reaction occupied, on the average, 50 per cent of his average deviation more than his average time. The same process was repeated for each individual, and the individual marks were averaged. This procedure, then, eliminates the absolute times, and also the absolute variabilities, and gives an average picture of the relative distribution of time throughout the list of twenty. The net result, on the average of the nine subjects, is as follows:

Position in list	1	2	3	4	5	6	7	8	9	10
Av. mark	+35	−50	−38	−22	−7	−17	−25	−32	−16	−10
P. E. of Av.	15	6	4	3	4	6	6	4	5	7

Position in list	11	12	13	14	15	16	17	18	19	20
Av. mark	+14	+26	+2	+34	+20	+32	+18	+6	+22	−4
P. E. of Av.	6	6	4	4	8	5	8	5	6	4

If the speed of reaction were uniform through the list of twenty, the average mark should be close to 0 throughout; but this is not the fact. In spite of the considerable variations and the rather large P.E., there can be no doubt that reactions 2-9 tend to be quicker than the average, and reactions 11-19 slower than the average. There is a slight slackening in the speed of reaction throughout the list. The first and last reactions are exceptions to this rule; for the first is slow, and the last is more rapid than those which immediately precede it. In regard to the first, individual differences are here very great and characteristic; and a fairly strong negative correlation (Pearson $r = -0.69$) appears between the time of an individual for the whole list and his relative time for the first word. This correlation is seen in the accompanying table.

Individual	Time for list of 20	Relative time for first reaction in per cent. of av. deviation
H	25.0	+179
Br	28.2	+ 41
Pf	29.0	+103
Pd	36.4	+ 10
R	37.4	+ 64
Wl	39.0	— 58
Bn	39.4	— 36
E	39.8	— 21
Wi	47.0	+ 32
Average	35.6	+ 35

The individual who is relatively slowest in the first reaction reacts more rapidly to stimuli in series than to single stimuli, and in this respect is rather exceptional. Probably he manages the "overlapping" of the successive acts better than most individuals. Accordingly his relative slowness in the first reaction may be probably explained as due to the necessary absence of overlapping at the start.

The fact, however, that *the reaction to the first word of a list is on the average slower than the reaction to an isolated word* shows that something besides overlapping and its absence are in question. Sometimes a subject reported that, in glancing at the beginning of a list, his eye had caught the second word along with the first, and that he was busied with the reaction to the second as early as with that to the first. It even happened, occasionally, that the reaction to the second word was ready before that to the first. This form of interference, incidental as it is to overlapping, would of course slacken the reaction to the first stimulus.

Two influences operate in reacting to a list that are absent in reacting to a single stimulus: interference and overlapping. The latter tends to accelerate the reactions, the former to slacken them, as compared with a reaction to an isolated simulus. Overlapping can not exert its accelerating effect upon the first reaction; and interference, also, would usually not become operative at the start, but special conditions, such as seeing the second word simultaneously with the first, may cause interference to be strongly evidenced at the very start.

Aside from the slowness and great variability of the first reaction, the most salient fact resulting from the above experiment is the quickness of reactions 2-10 as compared with reactions 11-19. Why should the speed decrease from the second reaction till near the close, and then increase again? "Fatigue" and "end-spurt" are the catch-words that readily occur to mind; but neither of them is specially explanatory. As for fatigue, so short a performance can hardly cause much fatigue of the genuine, metabolic sort. Interference seems to be a more probable conception. Each succeeding stimulus, and each reaction, tend to evoke associations that are of no service for the purpose of the test. These must be repressed; all their allurements brushed aside. A straight course must be steered in spite of many cross currents. As these deflecting tendencies continually accumulate with the addition of fresh stimuli and reactions, the likelihood of disturbance increases.

The increase in speed at the very close can probably be understood as incidental to overlapping; for, though overlapping leads on the whole to increase in speed, it does require, at every moment, a division of activity between two or more reactions. At the close, this division of activity ceases, and the last reaction receives the benefit of the overlapping without any of the incidental drawbacks such as were mentioned above in relation to the initial reaction.

The interest of the above experiment, in connection with the matter of tests, is the demonstration that the list test brings in factors—call them interference and overlapping, or call them fatigue, end-spurt, etc.—which are not present in reactions to isolated stimulus words. The list test reaches a more complicated mental performance and calls for a higher degree of control.

It was desired to see whether a shorter list would show the same time-curve as the list of twenty, and whether a list of ten words might not be essentially equivalent to ten separate stimuli. The experiment was of the same general character as above described, but was done in a rougher way. Instead of employing a rotating drum, the experimenter held the watch to his ear and with his pencil made wavy lines in time with the tick-

ing of the watch—the record so left resembling the trace of a tuning fork. At each reaction of the subject, the experimenter made a break in his time curve, and thus recorded the time of each single reaction. The accuracy of these times is about equal to that of ordinary stop watch readings. Meanwhile, by consulting his watch face at the end of the list, the experimenter had the time for the whole list. This double use of the watch can be recommended when list-tests are used, for the record of the single reaction, even if not highly accurate, is of value as showing how much of the time is lost in a few slow reactions, and as making possible the calculation of the median as well as the average time. Some practise is of course necessary before the experimenter can successfully use this device.

The results of this experiment were treated by the same statistical method as above described for the preceding experiment. Thirteen subjects served, each reacting to 13-24 lists, the total number of lists being 243. The average distribution of time in a list is shown in the following table, the explanation of which is the same as given on p. 15 for the preceding table.

Position in list	1	2	3	4	5	6	7	8	9	10
Av. mark	+179	−45	−33	−25	−11	−2	−19	−10	+3	−41
P. E. of Av.	12	7	4	6	3	7	7	4	5	7

It is quite possible that the very long relative time indicated for the first reaction is in part an artefact; but there is no doubt that the first reaction is slow, and that the last reaction is quicker than those that immediately precede it, just as was the case in the list of twenty. Further, there is a gradual increase of reaction time from the second to the ninth reaction. This slackening is less marked than in the list of twenty, but it is still present to a degree. It appears in the average results of 9 out of the 13 individuals; and the 4 who do not show a slackening show no progressive change in either direction. The conclusion is that the same factors are operative in the shorter as in the longer list, though not to as high a degree.

The results of the preceding experiments may well be compared with those of an experiment in which the stimulus words were

presented orally and separately. Five subjects were examined, each 16 times, with the same set of 20 words in different orders. The time-curve for these results presents an entirely different picture from that obtained in the preceding experiments. The first reaction is not slow, but on the average, one of the quickest; the last reaction is, on the average, one of the slowest, and yet there is no progressive slackening of the reactions from the first to the last, but the speed remains, on the whole, very uniform throughout. The difference in the time curve of the two modes of procedure is probably to be explained by reference to interferences: when the series of reactions is continuous, interferences tend to accumulate with the progress of the series, but when a brief interval of rest intervenes between the successive reactions the interferences tend to disappear. It may be concluded that the continuous reaction to a series of stimuli is a more complex process than the reaction to a single stimulus, and requires a higher grade of control. The two forms of test are not therefore equivalent, and each may be a good test; but, for a start, preference should be given to the simpler form, namely to the reaction to separate stimuli. Some associations, however, such as the naming of colors or other familiar objects, or the simplest arithmetical associations, are too rapid to be timed, singly, by the stop watch.

There is another advantage in the timing of single associative reactions over the timing of a series. The latter method gives indeed the average association time for the stimuli used, but (unless supplemented by some such device as employed above for getting the single times) it shows nothing of the distribution of the association times. In particular, since a series is likely to contain a few reactions much longer than the rest, the average time is apt to differ considerably from the median or the mode, and therefore not to be fully typical. The very slow reactions are usually due to rather special causes, and their great influence on the average is undesirable. The best procedure would seem to be that of timing the single reactions, and using the median as the typical measure.

Where a list-test, or continuous test, is employed, our experience leads us to favor a rather short list. A list of ten stimulus

words can be timed with sufficient accuracy, and it is freer, on the whole, from interferences of a disturbing character. In spite of this judgment in favor of shorter lists, we have presented lists of considerable length, partly to provide for use with a time limit—which is the procedure favored by Professor Thorndike, Dr. Whitley, and some other investigators of wide experience—and partly to provide a sufficient list of stimuli for separate reactions, when that is the method adopted. When the continuous method is used with an amount limit, it would be better to cut the lists in half and take two readings. Two short tests are better than one long one, because the average of the two is freer from the influence of momentary disturbances, and because it allows better for the effects of adaptation to the novel conditions of the test.

3. INSTRUCTIONS TO THE SUBJECT.

The necessity of uniform instructions has often been insisted upon, and sometimes the instructions have been reduced to a set formula, in order that all individuals tested, receiving the same instructions, may be treated alike. A set formula is, however, no guarantee that the subjects are treated alike, for some may not comprehend the formula. With a rigid form of instructions, the test becomes partly one of the individual's ability to understand the instructions, and only partly a test of the function exercised by the test material. It would be better to provide separate tests for ability to understand instructions, and eliminate this factor from other tests, so as to make each, as far as possible, a test of one function. Proper comprehension of the experiment by the subject must not be sacrificed to an ideal uniformity of instruction. It matters little by what method is attained the uniform *result* of the understanding of the experimental task. However, there can be little question that the best method for this result is that of "learning by doing," and that the subject should learn and demonstrate his capacity for the prescribed reactions by going through them. Instruction should proceed by description, illustration, and execution. The subject should first be told clearly the nature of the test; then if possible he should see

the operator perform some example of it, and finally he should execute samples of new material himself. None of the preliminary samples should duplicate the actual test material; nor should the preliminary trials be multiplied beyond what is necessary to insure the understanding of the test and the first strain of adaptation to it. As a rule it may be wise to allow the subject to make correct reactions to two samples before passing to the actual test.

Samples for use in instructing the subject should be prepared beforehand and lie ready to the experimenter's hand. For his convenience, it has seemed best to us to provide blanks containing samples of most of the tests—one blank containing samples for several tests.

The number-checking tests and the directions test—to be described later—require not oral but written responses, and their execution uses up the blank provided for each subject, and also the sample used by each subject. Since this is not the case with the remaining experiments, the instructional material of these two tests is best kept separate from that for the remainder. It is reproduced[1] on the accompanying page ("Instructional Material I"). For the two forms of the number-checking test, two lines of material, organized in the same way as for the actual test blank, are provided. For the directions test, there are three directions similar to those on its actual test blank. For the remainder of the tests, the same blank may be used for instructing an indefinite number of subjects ("Instructional Material II"). For each of the forms of addition test, successions of seven figures are supplied, the subject after verbal instruction and illustration reacting to these precisely as to the subsequent test material. There are three sample words for the opposites test, and also three samples each for the considerable number of tests of partially controlled association. The substitution test and the color naming test supply their own instructional material. The six words provided for the free association test are contained neither in the Kent-Rosanoff experiment nor in the supplementary thousand-word list.

[1] Here, and in some other cases later on, the exact type, etc., of the blank is not reproduced.

Instructional Material I.

Number Checking Test.

Form A. 458792360174186059235960842317821307564945827639017608439512194725083636457012986528394017237694185 0

Form B.

215864	381592	826739	967814	371245	942861
876395	269517	712983	368459	326748	258647

Directions test:

1. Write any number larger than 16.
2. Add one more dot to the largest group . ... ..
3. Put a cross over the angle that opens downwards ∨ ∧

Instructional Material II.

Addition Test.

Kraepelin Form	*Constant Increment Form*	*Opposites Test*
4	32	better
9	47	
3	21	glad
8	53	
6	39	straight
5	28	
2	65	

Logical Relation Tests.

vb-obj	*supraord*	*subord*	*pt-wh*
cut	horse	flower	roof
buy	Paris	lake	tail
bend	potato	game	Germany

wh-pt	*agt-act*	*act-agt*	*att-subst*
wheel	train	shines	cold
Europe	frog	howls	cheap
brush	sun	crawls	narrow

Mixed Relations Test.

Box—square	Orange—
Woman—husband	Man—
East—west	Day—
Penny—copper	Nail—
Asia—China	Europe—
Grain—sand	Drop—
Am—was	Have—

Free Association Test.

fox	cure
apple	quick
fork	grass

Even after thorough and apparently successful instructions, it will occasionally happen that the subject's reaction to the actual test material shows at once that he is on the wrong track. In such cases the test must be called off. But if the experimenter is provided with pairs of equivalent tests—as can always be accomplished in the present series, either from the provision of duplicate test blanks, or from the cutting of a long blank into two, as previously (p. 20) recommended—then the material of the first blank may be used for further instructional samples, and the actual test carried out with the equivalent blank.

There is some difficulty in bringing all subjects to an equality in their attitude towards the matter of speed. Occasionally it happens that an individual does not try for speed, but only for accuracy or distinction or even for an introspective study of his performance. Since the time of the performance is the important matter in utilizing the results, it is unfortunate when a good subject fails to make an effort for speed. We have seen the standing of an individual among his fellows completely changed, in the middle of a series of tests, by his being informed that his performance was slower than the average; his times were at once cut nearly in half, while his accuracy was not lessened. We recommend that the instructions include some such statement as the following: "The main thing that we are after is to see how rapidly your mind can act. You need not be afraid to put on speed, for the test is easy and you are not likely to make mistakes. Of course, you should keep on the right track and not make mistakes, and for every mistake you will be docked a little—about two per cent" (or one per cent if the whole of long blanks is used). The docking by two per cent is, of course, purely arbitrary; and it may be desirable with some classes of subjects to make a larger correction for errors; but our experience with these simple tests has not revealed the need of any corrections at all for errors. It is probably wise, however, to mention the possibility of errors at the beginning, and to have an understanding, at least with mature subjects, as to the degree of importance attached to them.

III. CANCELLATION TESTS—THE NUMBER-CHECKING TEST

Controlled association tests are, in method, somewhat analogous to choice reaction experiments. Like them, they may involve a certain response to every stimulus, according to a prearranged scheme of reaction (the B-method of Wundt), or they may involve one single reaction to a single kind of stimulus among a heterogeneous group of stimuli, the C-method of Wundt. Here the subject either reacts or does not, i. e. has the choice between movement and rest. The present test is the only representative of this method among the experiments to be described. The general idea is the presentation to the subject of a blank upon which are printed a large number of different letters, figures or designs, and requiring the recognition of each of a certain symbol to be indicated by marking. There are measured the speed and accuracy with which this is done. This experimental conception is not a novel one, having had its origin something over 15 years ago, and having, in various forms, played a part in a considerable number of subsequent investigations. In company with the diverse forms of material that are available for its performance, it has also borne a diversity of names; the most familiar single form is probably the so-called A-test, first mentioned by Cattell and Farrand, which has long played a part in the Columbia Freshman Tests as a measure of "rate of perception." As a rule, it seems desirable to know a test rather by an individually descriptive title than by an at best somewhat vaguely defined mental function with which it may be related. We thus offer the present attempt at the standardization of this method under the name of the "Number-checking Test;" prepared for special purposes in two forms.

Form A. As this is, so far as we know, the first time that a form made up exclusively of numerals has been offered for this purpose, it may not be amiss to give some account of the considerations which led to their adoption. The use of ordinary con-

textual material has its advantage in certain individual applications of the test, but is out of place in standard form, because the content of the text employed disturbs the attention of the subject to the experimental task, and disturbs it in different ways and in different degrees for different subjects, thereby violating the first principles upon which these tests are constructed. All this quite apart from its absolutely chaotic arrangement of the significant stimuli. Pied type obviates most of these difficulties, especially if methodically arranged; yet even here it is difficult to avoid vocable or other combinations of character that would have significance external to the test. Geometrical forms satisfy the conditions better, but here arises the difficulty of finding a proper number of distinct geometrical forms, small enough for the requisite purpose, and recognizable with sufficient readiness; for every effort must be made to obviate false reactions. These considerations seemed to point pretty definitely toward the use of Arabic numerals. They are as readily recognizable as the letters, and the chance of any specially suggestive collocation is infinitesimal in comparison with the letters.. The number of the symbols, 10, lends itself logically to the use of 100 of each symbol in a blank of 1000 symbols, which is as long as there is any necessity of making such a blank.

The general character of the blank being thus determined, the arrangement of the material took place as follows: Twenty lines of fifty symbols each, properly spaced and justified, provided when printed in the regular eight point type a printed space with suitable margins on the regular blank sheets of the experiments. The arrangement of each line is such that it contains five each of the symbols 1,2,3,4,5,6,7,8,9,0. Each successive fifth of the line (10 symbols) contains each of the ten different symbols once. This arrangement of itself obviates the occurrence of "runs" of two or more of the same symbol, except at the beginning and end of each ten, and it was not allowed to occur here. The first ten lines of the blank being completed in this way, the second ten lines were constructed by reversing this arrangement. This procedure assured the approximate equality of the two halves of the experiment as well as the uniformity in the distribution of all characters throughout.

5168492370127504869341890256371756089243786904312 5
7805134269240976153383204157968481267390537908652 14
3597846102518237496085932410769345208617933167584 02
2739650814973615028404785962130931566487242356790 81
4253017986386091547293674801256493120758614274905 38
9470385621609382714578109643525279416380104823795 56
0982561743833546920176021379584218495703609521867 43
1046279538062843975127456038913507842169568192438 70
8614923057451728630919568827408620735491247350186 9
6321708495794150382656832718409706839152485043126 97
7962134058425193860790481723656628305149759480712 36
9681053742194537026804723865919036827154750329416 8
0734291865961248705319830654721579348260835972640 1
3476812590630759481248597812067102964533834716526 90
6597328401083614972525346901875417283906126583074 9
8350947216857021394652108476392745190683689710352 4
1809765324278465139031269587404820516379418056693 72
2048576139716802543967014239580694732815201648795 3
4125680973509376218486975140238351679042962481506 87
5213409687342980657173652098143968405721073294861 5

Special researches may call for different blanks prepared on the same principle as those submitted. A great number of blanks may be derived from the present one by methodically replacing each figure by some other, thus everywhere substituting 1 for 0, 2 for 1, etc. The test has many desirable features as a measure of fluctuations in continued work. Where it is used for this purpose a much longer blank would probably be desirable, in which the subject should check a greater variety of numbers, say all the odd numbers. The time for every line could be noted, or if greater precision were required, the time could be reported electrically by the graphic method.[1] The experiment should also be especially convenient in studies of interference, requiring a subject practised for one number to check different numbers, and the like.

With the above arrangement, similar as it is for all symbols, it makes no difference, so far as the arrangement is concerned, which one of the ten symbols the subject is instructed to check. The symbols do not however appear to be equally easy and their order in this respect should follow their order of distinctiveness; so far as is determined, 1 and 7 seem to be the easiest, next 0 and 4, next 2, 3, 5, 8, and the hardest 6, 9. In the present experiments the symbol 0 has uniformly been the one checked. Since some subjects may discover essential features in the arrangement of the blank while others do not, and also with a view to obviating omissions, the subject should be uniformly instructed that there are in each line five of the symbol they are to check. This reduces omissions to an insignificant minimum, if it does not obviate them altogether; and this advantage probably more than compensates for the occasional delays that result from missing the next symbol in succession and having to go back and search for it. In practise the procedure recommended seems the more desirable with the subjects so far employed.

After the subject has shown proper understanding of the experiment by correctly executing the two similarly arranged lines of the instructional material, he begins the execution of the test

[1] Cf. von Voss, Ueber die Schwankungen der geistigen Arbeitsleistung, *Ps. Arb.*, II, p. 300.

blank according to the directions already outlined. A pencil should be used for checking, and while it were easy to appear hypercritical in this respect, even in so minor detail as the character of the pencil, approximate uniformity should be sought for among all subjects between whom direct comparison is to be made. It should be one of moderate softness, not harder than the No. 2 grade, otherwise the markings may require more than the accustomed writing effort to make them properly distinct. The timing of the complete execution of the blank is by the stop-watch, and it is advisable to take the time for the halves also; unless the watch is provided with a split-stop, it is scarcely feasible to take the half-time closer than the nearest second.

Probably no other of the present tests shows such irrelevant differences in the manner in which different subjects execute it. The check-marks vary from light, almost haphazard strokes, to heavy, labored scorings. It is impossible, even were it desirable, to secure uniformity in these respects. In such matters the subject must adopt his own optimum method. While it is not clear that any special advantage accrues from this procedure, the same must apply to the habit of some subjects to check alternate lines backward. More significant data are probably obtained by merely noting whether the subject modifies the more natural behavior to the alternating directions, and if this is done immediately. A few subjects, having checked the two lines of the instructions blank, have shown a tendency to check only the first two lines of the experimental blank, and subjects should clearly understand that the *entire* blank is to be gone through. Half of the blank is however long enough.

On rare occasions a wrong symbol is checked; this seems to be more frequent in the later stages of practise. As with the omissions, the influence of such errors is due less to their *direct* distortion of the final time than to the fact that the subject is quite apt to observe them and be disturbed by their occurrence. On the basis of false reactions, the test has not shown, in the writers' hands, workable individual differences in the "accuracy" of permance. There are perfectly distinct differences in the time of performance, and there seems little reason, in the present test, for extending the scoring beyond this single factor.

In this respect, the range of normal performance in the test at the beginning of practise would seem to lie between 100 and 200 seconds, with an average of 133 for the whole blank. The subjects averaging this figure were 20 men and 20 women of a similar group. It is worthy of note that the women averaged distinctly faster than the men, 123 as against 145 seconds, and also somewhat less variable, the m. v.'s approximating 15 and 22. One of the writers has usually found the sex difference in variability in the other direction, and, as is brought out just below, the difference between this and the succeeding experiment was somewhat more marked in the women. In both groups the second half of the experiment averages slightly faster than the first half, though there are great variations in this respect.

The function with which the test is concerned is one rather susceptible to practise, as anyone may readily discover for himself. In a succeeding experiment the average times dropped to 116 and 134 for the women and men respectively, and the first half of the experiment averages a little faster in the women. Much of the significance of the experiment depends on the preservation of the individual relationships originally indicated. They are better preserved in the men, the orders corresponding within 13% for them as against 24% for the women, though this is largely due to two cases, one of whom rises fourteen places, the other dropping ten. The function is evidently one whose expression in the test can be distorted by incidental factors that are as yet very imperfectly understood, and the advisability of the greatest possible standardization of the experimental conditions and material is only further emphasized.

For subjects who at the outset evince a fairly definite quality of performance, practise does not tend to any great alteration of relative position. There seems to be less individual difference in susceptibility to practise than in the range of performance at the beginning of practise.

Form B. This variant is termed the "number-group checking test." It will be borne in mind that the conditions of the above discussed experiment have been frequently altered to require the checking of more than one symbol. The most familiar form of

this is the so-called *a-t test*, in which the subject, using a passage of connected prose, checks every word containing both the letters *a* and *t*. Other tests of a similar nature have been known by the names of the symbols checked. This is not the same as merely requiring the subject to check all of two or even more symbols, since the symbols have now to occur in a definite combination. The original blank was therefore not adapted to this purpose, and in order to meet in an analogous way the possible requirements of such a test, a special form was constructed for it.

This test contains all the combinations of nine digits taken six at a time. The number of such combinations $= 9.8.7.6.5.4/1.2.3.4.5.6. = 84$. Each single digit occurs ${}_8C_5 = 56$ times; each pair of digits ${}_7C_4 = 35$ times; each three, ${}_6C_3 = 20$ times; each four, ${}_5C_2 = 10$ times; and each five, ${}_4C_1 = 4$ times.

In preparing the test, a separate card was used for each of the 84 combinations, which were first written in the order of the digits, e. g., 134689, 125678, 245789, etc. Then all the 35 cards containing both of the digits 1 and 2 were separated out of the pack, and this pair of digits was assigned to the various possible positions in the group of six with approximately equal frequency. At the end of this operation, therefore, there were, as nearly as possible, equal numbers of groups arranged in the following ways (the dots indicating positions not yet filled):

12....	1.2...	1..2..	1...2.	1....2
.12...	.1.2..	.1..2.	.1...2	
..12..	..1.2.	..1..2		
...12.	...1.2			
....12				
21....	2.1...	2..1..	2...1.	2....1
.21...	.2.1..	.2..1.	.2...1	
..21..	..2.1.	..2..1		
...21.	...2.1			
....21				

Owing to the preliminary shuffling of the cards, the relation of these positions of the pair 1,2 to the remaining composition of the combinations was haphazard.

Next all of the cards containing the pair 1,3 were separated

out of the pack and shuffled, and the assignment of the pair 1,3 to its several possible positions with equal frequency was undertaken in the same manner as before. The operation of pure chance was somewhat limited by the previous assignment of the digits 1 and 2. The same operation was repeated with each pair of digits, 14, 15, 16, 17, 18, 19, 23, 24, 25, 26, 27, 28, 29, 34, 35, 36, 37, 38, 39, 45, 46, 47, 48, 49, 56, 57, 58, 59, 67, 68, 69, 78, 79, 89, in this order. As a matter of course, the operation of chance became more and more restricted as this work progressed; and it became more and more difficult and finally impossible to insure that the several positions of a pair should occur with equal frequency. All that could be hoped for was approximate equality of arrangement of the different pairs; and a review of the result, at the close of this operation, showed that approximate equality had been attained. Each pair of digits, therefore, will be found not far from once or twice in each of its 30 possible positions. Since it is believed that the chief application of this blank will be for checking the groups containing a given pair of digits, this approximate equality of arrangement of all the possible pairs is probably what is most needed. Some care was however taken also to avoid undue repetitions of the same arrangements of groups of three digits.

After the internal arrangement of each combination of six digits had thus been determined, the cards were again thoroughly shuffled in order to determine a chance order of the combinations. But the order was not left entirely to chance, for the immediate recurrence of the same pair of digits in the same position was avoided. Such immediate recurrences are likely to be noticed and remembered and so interfere with the repeated use of the blank with the same pair.

The result of all the operations so far was the obtaining of a series of the 84 combinations of the nine digits, taken six at a time, with approximate equality in position of each pair of digits, and approximately chance order of the combinations. Now it was desired to double the length of the blank, and in such a way that the second half of the blank should be equivalent to the first half. Each of the 84 combinations was therefore to be repeated in a

new permutation, and the equality in the arrangement of pairs was to be maintained. This could be accomplished, without need of going through all the operations involved in preparing the first half of the blank, by simply permuting each of the groups of the first half in the same way. Each of them was, in fact, permuted according to the following scheme: 2, 4, 6, 1, 3, 5. The second set of 84 groups thus obtained must have all the characters impressed on the original set, only with change of the absolute digits; and there can be no duplication between the two sets. By similar schemes of permutation, 720 sets in all could be obtained without duplication at any point.

After the above blank had been printed, it was checked up in every way to see whether it was according to specifications, and found to be correct. Since, in spite of the statistical equality between the halves, there might still be inequalities in practise, due to the element of chance entering into the arrangement, the halves were empirically compared, by taking the time for each half both for each single digit and for each pair. Our results are not, indeed, numerous enough to establish the precise equality of the halves, but they give no reason to suppose the halves different in any respect.

Were the nine digits of equal perceptibility, this blank would afford a large number of equivalent tests. But the digits are quite unequal in perceptibility; and it is therefore necessary to establish the relative difficulty of the several tests by trial. We have tried the following tests: (1) cancelling the groups containing each single digit, 9 different tests; (2) cancelling the groups containing each pair of digits, 36 different tests; (3) a few of the 84 possible tests in cancelling groups containing three assigned digits. The results, though not as extensive as could be wished, show much regularity and can probably be taken as indicating, approximately, the relative difficulty of the several tests.

(1) Cancelling of groups containing a specified single digit. In this, as in all the following results, one half of the blank was used at a time, and the time is given in seconds.

Seven subjects, previously untrained in this test, were tested

983642	168379	694517	253914	745682	158923	729648
426357	372159	754936	297835	627519	786531	731469
654173	947386	589761	134852	146237	194526	936425
837162	691324	814536	326175	368792	549826	572194
458671	971648	479612	495683	784295	817243	916328
275148	318495	635728	596873	982563	431289	381647
513978	182765	615832	851279	498136	356719	412789
197584	563792	748315	861395	421856	973124	125437
918654	846975	453867	281463	213956	651274	526987
397841	961872	248691	574389	532416	723964	473519
872351	327984	437528	864712	825916	682543	534169
923871	632791	765429	235849	672834	295481	349257
867314	462758	486592	198537	871596	164985	247153
963458	981374	156843	259671	762491	983567	579361
345962	941258	182653	561487	435781	179428	731825
672389	346521	427163	281937	672539	985273	956142
312876	853926	587436	296851	784623	875126	513647
934612	739548	843216	215367	916483	294378	768914
954178	371629	529817	436978	123874	957641	682917
719325	294736	639187	286415	593182	297568	145389
594231	389254	196235	825749	461289	378652	672841
349716	427395	138962	268794	524617	358472	319546
714932	759431	382145	853624	714529	635819	237465
649752	718254	596743	862934	851763	329418	495867

in cancelling each of the nine digits, the order of these nine tests being different with different subjects, so that any transferred practise effect from one digit to another is to a large extent equalized in the average of the seven subjects.

Digits	1	2	3	4	5	6	7	8	9
Av. time	43.5	63.0	59.9	53.7	61.4	70.9	54.2	57.4	65.1
A. D.	4.0	7.9	7.1	5.4	6.3	6.2	5.5	7.6	11.0
P. E.	1.3	2.6	2.3	1.7	2.0	2.0	1.8	2.4	3.4
Total range	37–56	53–81	46–74	42–63	48–78	62–82	39–67	43–69	50–78

One subject made eight trials with each digit, showing rather slight improvement after the second round. For trials 3-8, his times average as follows:

Digit	1	2	3	4	5	6	7	8	9
Av.	31.0	45.5	40.1	38.7	42.9	50.8	36.1	37.6	44.4
A. D.	2.2	2.2	1.1	1.9	1.1	1.0	1.3	1.7	0.5
P. E.	0.7	0.7	0.4	0.6	0.4	0.3	0.4	0.6	0.2

Much the easiest digit to cancel is 1. It is easiest with every subject tested, and in every trial. If the times for the other digits are expressed as per cents of the time for the digit 1, the following are the relative times:

Digit	1	2	3	4	5	6	7	8	9
Relative time, av. of 7 unpractised subjects	100	145	138	123	141	163	125	132	150
Relative time, one practised subject...	100	147	130	125	138	164	116	121	143
Relative time, 3 subjects, 2nd trial....	100	156	131	129	130	163	121	125	141
Av. of above, with double weight allowed for first trial	100	148	134	125	138	163	122	128	146

The relative times in the three sets of results are in fairly close agreement, and the combination in the last line of the table can certainly be relied on within a few per cent. It is certain that 6 is the hardest digit to find, as 1 is the easiest. The important practical question is whether any digits are of nearly

equal difficulty, so as to be available for equivalent tests. Approximate equivalence is assured for the following pairs:

4 and 7
3 and 5
2 and 9

Further experience with the tests will probably show the need for slight corrections in treating these pairs as equivalent.

When only one pair of equivalent tests is desired, the easiest is probably the best, especially as our results show that errors and omissions are less frequent with the digits that give shortest times. Thus, the seven unpractised subjects whose times are reported above gave the following average number of errors (mostly omissions) per test:

In cancelling for the digit 1	0	errors
In cancelling for the digit 4, 7 or 8	½	"
In cancelling for the digit 3 or 5	1	"
In cancelling for the digit 2 or 9	1½	"
In cancelling for the digit 6	2	"

Since the time measure is of most value when errors are absent, the digit 1 is indicated as the best to use, except when there is need of an equivalent pair of tests; in that case, 4 and 7 are the best to use.

In regard to a correction for errors, our experience has not shown the need of one. Our subjects have not seemed to save time by omissions, but the time has been about the same either with no errors or with one or two or even three omissions. These subjects were, to be sure, serious and attentive; and it is likely that a more varied experience with the test would show the desirability of correcting for errors. We judge that the corrections should be small, and suggest the addition of 2 per cent. of the subject's time as penalty for each error or omission,[1] when one half of the blank is used; or 1 per cent. when the whole blank is used.

(2) Cancelling of groups containing a specified pair of

[1] An expeditious method of detecting errors is afforded by a key on transparent paper, to be laid over the blank. Whichever digit is used, the number of groups to be checked is 56 in each half of the blank.

digits. One subject has made six trials of each of the 36 tests of this sort; and another subject has made one trial of each. The results appear to have sufficient regularity to indicate the relative difficulty of the several pairs, and to show something regarding the mental process involved in this form of test.

The time occupied in checking a pair of digits is always longer than the time for checking either digit alone, but less than the sum of the times for checking the digits separately. For example, a subject takes 42 secs. to cancel the digit 4, and 48 secs. to cancel 9; to cancel groups containing both 4 and 9 takes him 64 secs., which is 71 per cent. of the sum of 42 and 48. The time for a pair is closely correlated with the sum of the times for the digits of the pair, and is usually equal to about 70 per cent. of this sum. The results are condensed into the following table.

RATIO, IN PER CENTS, OF THE TIME FOR CHECKING A PAIR OF DIGITS TO THE SUM OF THE TIMES FOR CHECKING THE DIGITS SEPARATELY

Subject	Average	P. E.	A. D.	Range
J. W. T.	67.3	0.4	3.1	60-75
R. S. W. first trial	72.2	0.6	4.5	62-82
R. S. W. after practise	72.5	0.4	3.0	66-79

Since this "ratio" is fairly uniform, it can be used, in connection with our previous table of times for checking single digits, to indicate the approximate times for checking pairs. Equivalent tests can be selected in this way; among the tests which appear from all our present results to be nearly equivalent, we recommend the following two pairs: 23 and 89. These have the advantage of not conflicting with the digits 4 and 7 recommended for use when single digits are to be cancelled. The time for either pair is about twice that for the single digit 1, or about one-and-a-half times that for the single digit 4 or 7.

It is possible, from comparison of the results of the two subjects in the above table, that there are genuine individual differences in the "ratio," i. e., in the speed of cancelling pairs as compared with the speed of cancelling single digits. Such differences may however, be merely the result of the relative degree of practise in the two sorts of test. If the subject has gone further in

his practise with pairs than in his practise with single digits, the ratio will evidently be small. If he has had some practise with single digits, but none with pairs, his first experience with a pair is likely to give a high ratio. Thus, subject R.S.W., after making two trials with each of the single digits, proceeded to try in succession each of the 36 pairs.

The average "ratio," for the successive quarters of this series, was as follows:

	Av.	A. D	P. E.
First quarter	76.5	3.4	1.0
Second quarter	71.9	4.3	1.2
Third quarter	72.6	2.8	0.8
Fourth quarter	67.9	3.2	0.9

He then resumed practise with the single digits, and afterwards returned to the pairs, obtaining then the average ratio of 72.5, as shown above. If practise is continued *pari passu* with single digits and with pairs, the ratio would probably remain in the neighborhood of 70 per cent. But at the very start, the time for a pair is likely to be about 77 per cent. of the sum of the times for the single digits; this is indicated also by less complete results from several other subjects.

The dependence of the "ratio" on practise has a bearing on the theory of this test. The fact that the ratio is high at the first experience in cancelling for a pair of digits shows that the detection of a pair of digits in a group is a specialized performance, not reducible to the acts of detecting the single digits. The detection of any specified pair of digits is no doubt a specialized performance, susceptible of very special training; this has indeed been shown in similar cases by Thorndike and Woodworth. But in the present series of tests, the pair of digits cancelled was changed with each new trial, so that the training visible in the lowering of the ratio from 77 per cent. to 72 or 70 per cent. is an example of transferred practise, and indicates that there is some element of skill common to the checking of all the pairs of digits.

Though the ratio varies within rather narrow limits and shows a comparatively small A.D. (as seen in the table on p. 36),

yet there is sufficient variation to make it possible that the ratio varies according to the nature, difficulty, etc., of the pair employed in the particular test. We have been unable to find any characteristic difference, however, according to the difficulty of the digits entering into the pair, except that, in case of subject R.S.W., the ratio is low for pairs containing the digit 1. The subject, after practise, gave the following ratio for pairs containing the several digits.

Pairs containing the digit	Average ratio	P. E.	A. D.
1	68.2	0.5	1.7
2	73.1	1.1	3.6
3	72.8	0.8	2.7
4	72.8	0.8	2.8
5	72.1	0.7	2.3
6	73.4	0.7	2.4
7	73.7	0.7	2.3
8	71.8	0.9	3.1
9	74.3	0.8	2.6

In general, the ratio seems not to depend on the digit; and the same negative result appears in case of the other subject, J. W. T., who moreover does not show anything characteristic of the digit 1. But the above results from subject R.S.W. show an unmistakeably lower ratio for pairs containing the digit 1. The distribution is fairly bimodal, the pairs containing 1 forming a group by themselves.

Some explanation of the low ratio for pairs containing the digit 1 is afforded by R.S.W.'s introspective account. It early occurred to him that a good device for cancelling groups containing a pair would be to look first for the easier digit of the pair, and thus to look for the harder digit only in the groups where the easier digit appeared. In practise, however, this device did not seem to him to work very well, except when the easier digit was 1; he tried to use the device also when the easier digit was 4, 7 or 8, but without subjective indications of success. When one of the digits was 1, groups containing it could be recognized in indirect vision, and thus many groups could be passed over altogether in direct vision. Subjectively, this method

of working required more effort but appeared successful. The objective records, as crystallized in the "ratio," show that the device was a success in the case of the digit 1.

Further consideration of this point may throw some light on the mental process involved in this test. If finding a pair were the same thing as finding the members of the pair, with no overlapping, the time for the pair would be the sum of the times for the digits composing the pair—instead of being, on the average, only 70 per cent of that sum. There must therefore be considerable overlapping or condensation. On the motor side, there is a possible condensation of the checking movement, but this is so quick and automatic anyway that abbreviating it has probably little to do with the shortening of the time. More strain is probably put on the eye movements when the speed of the work approaches its maximum (about 3 groups covered per second); but since this maximum is not approached, in our results so far, except by one subject in case of the single digit 1, the probability is that the demands made on the eye are well within its motor capacity. The difficulty of these tests is mainly perceptional, and the overlapping which is effective in finding pairs of digits must occur in the perceptive process.

If the device described above as adopted by one subject in finding pairs of digits—a device which has frequently been adopted by other subjects in similar tests—if this device represented the essentials of finding the pair quickly, then the following calculation should hold good. The subject looks first for only one digit, and where he finds it looks for the other one. The task of looking for the second digit would be necessary only in ⅔ of the total number of groups in the blank (since the first digit, or any digit, is present in 56 out of the 84 groups). If therefore this plan were carried out systematically and without hitch, the time for checking a pair should be equal to the time for checking the first digit plus ⅔ of the time necessary to check the second digit in the entire blank. For example, in checking the groups containing both 1 and 2, the time would be that needed to find the 1's (and this is 31 secs.) plus ⅔ of the time necessary to go through the blank for 2, namely ⅔ of 45.5 seconds, or 30.3 seconds; which added to

31.0 secs. gives 61.3 secs. as the calculated time for checking the pair 12. But the observed time is considerably less than this, namely 53 secs.; and this same discrepancy between the calculated and observed values obtains in every instance. The time for checking a pair is never as long as it would be if the above device were followed systematically.

This device does not therefore constitute the essential mechanism of checking a pair of digits. The device seemed to work well with one subject, in case one of the digits was 1; but its conscious use only reduced the "ratio" from 72.5 to 68.2, or 4 units, whereas something else reduced the average pair from the maximum of 100 to 72.5, or 27 points. There must accordingly be some form of overlapping of which the subject is not clearly conscious, but which is much more efficacious than the best devices which he consciously adopts. Introspection gives some hints at such an overlapping. Sometimes, indeed, a group is successively examined for the two digits composing the pair; but this is rather the exception. Often the presence of both digits is simultaneously recognized; and still oftener the absence of the pair is recognized without a clear consciousness of which of the two digits is lacking.

(3) Cancelling of groups containing three specified digits. Our data here are limited to 25 tests with one subject. The time occupied in this test is, again, closely correlated with the sum of the times for the three component digits, and is equal to about 70 per cent. of the sum of these times. Apparently the ratio is slightly lower for three digits than for a pair, for the subject, R.S.W., gives an average ratio of 68.3 per cent., with A.D. of 5.4, and P.E. of the average of 0.9. This average is thus probably lower than the average of 72.5 obtained by this subject with pairs.

When 1 is one of the three digits, the average ratio is 62.0, A.D. being 3.4 and P.E. 1.1. The subject adopted the same device as in pairs containing the digit 1, and the results here are confirmatory of what has previously been said.

For two approximately equivalent tests, we recommend cancelling for 146 and for 257. The time for each of these is about

2.5-3 times that for the single digit 1, or about 2-2.5 times that for the single digit 4 or 7.

Use of the number-checking blanks with laboratory classes. As suggested above (p. 27), the number-checking blanks could readily be adapted for experiments in continued mental work, interference, etc. One of the writers has used Form B, the "Number-group Blank," with success in an experiment in practise and "transfer." Half of the blank being used as the unit, the subject first checked the groups containing the digit 6, then took a practise series of ten units with the digit 7; then one unit again with 6; then another practise series with digit 4; then one unit with 7 and finally one more with 6. The two methods of studying "transfer," namely the "cross-section" method, and what may be called the "successive practise curve" method,[1] are combined in this experiment. The tests with digit 6 give cross-sections before and after practise with other digits; and since the digits 4 and 7 are equally hard to find, the practise curve with the one, following that with the other, should show the effects of the preceding practise. Transfer is pretty sure to be in evidence in each student's results; these need, to be sure, some correction from control experiments in which the cross-sections are taken without the intervening practice.[2]

[1] Introduced by Bair, "The Practice Curve", Psychol. Rev., Monogr. Suppl. No. 19, 1902.

[2] See W. F. Dearborn, *Psychol. Bulletin*, 1909, 6, 44.

IV. ADDITION TESTS

For rigidity of associative control, no experiments surpass those involving the simple arithmetical processes. In these a certain arithmetical task is visually presented to the subject, and efficiency is measured in terms of time and error. To provide an objective criterion of the performance the subject is required to speak or write the result. According to what has gone before, oral response is employed exclusively in the present experiments. One advantage of the written response is thus dispensed with, namely the permanency of the record, through which to check its accuracy. This difficulty is best obviated through providing the operator with a key upon which the correct reactions are noted. The operator follows the responses of the subject on the key and so keeps account of the data to be recorded.

Such experiments with arithmetical processes have an almost infinite range of difficulty, varying in practise from the simple addition of a pair of digits to the mental multiplication of three and even four place numbers. The chief advantage of the former is their freedom from errors; of the latter, the greater proportion of time spent in the essential work of the test. At first glance, one might consider that this same consideration, which leads to the substitution of oral for written response, should lead to the rejection of the easier and adoption of the more difficult experimental material. But it were very easy to press this advantage too far, especially in tests that are intended for anything like general employment. The more complex intellectual associations would result in the average individual in an impossible number of errors, if indeed they did not prove too much for his patience as well as his powers. A test not intended for limited application should not be one limiting the subjects who can respond to it; the tests to be described here, therefore, deal with the simpler arithmetical processes, regularly of addition, though the material prepared is adaptable in various ways.

1. The particular form of addition test with which the most

work has been done, and whose properties with reference to the work curve are best understood, is that of the Kraepelinian *Rechenhefte.* This is a pamphlet of twenty-four pages, upon each of which are printed nine vertical columns of 32 single digits in apparently random succession. It is possible to experiment with this material by continuously adding the successive digits, and announcing the sum total at stated points. The disadvantages of this procedure are very numerous and do not call for consideration here. A decidedly preferable method is the simple addition of the successive pairs of digits. That is, the first four figures in the *Rechenheft* being 8, 3, 5, 7 the sums announced by the subject are 11, 8 and 12. The subject continues to announce the sum of every figure plus the one next below it. Precise control of the whole process, both as to accuracy and time, is thus secured. As before mentioned, the operator checks the correctness of the sums, notes errors, and the amount performed within specified times. The usual periods of work with the Kraepelinian test have been of five and ten minutes each, and it has also been customary to record the amount of work done during the single minutes. The subject should not, as has been done, be called upon to make the records; all such tasks devolve properly upon the experimenter. It does not appear that a significant portion of the time is consumed in the motor process of response. At the beginning of practise, the number of additions made in five minutes is usually under two hundred. It need scarcely be said that the sums themselves could be read in a much shorter time; maximum speed of reading aloud in normal individuals averages not far from 100 words in 30 seconds.

Used in the above way the Kraepelin *Rechenheft* contains 31 additions per column, 279 per page. While the unpractised subject is not likely to do more than this in five minutes, a little practise will soon take him over the page, and it may be considered always advisable to open the *Rechenheft* to two full pages, (it should be held open with a clip, amply providing for any common performance without turning a leaf. One of our subjects, after prolonged practise, occasionally reached a figure above 558 additions in the five minute period. There is uniformly a con-

siderable practise effect in the test, perhaps due partly to the unfamiliar sort of reaction required to the material, but, as with the number-checking test, the individual differences in susceptibility to practise are small in comparison to the differences in the amount of work performed.

The Kraepelin *Rechenheft* is practically the only form of this experiment that permits long continued tests of the same subject with sufficient uniformity of experimental material; but as the present tests are not designed as practise experiments, it was thought advisable to construct a more convenient blank for a single or small number of determinations. It was also endeavored to improve on certain minor features of the *Rechenhefte,* as the odd number of columns and additions in each column.

The blank, reproduced on the page opposite, contains 24 columns of 26 digits (25 additions),—in all, 600 additions. This is probably beyond the 5-minute capacity of the normal subject at the limit of practise. The columns are arranged in six groups of four, each thus containing 100 additions with 104 digits. The scheme of the distribution of the digits was simpler than in the number-checking test. 99 slips were prepared, eleven for each of the nine digits. The five remaining slips bore the digit 5. Random drawings were made from the group of slips, and the digits were written in the columns in the order in which they were drawn. Each set of four columns (100 additions) is therefore a unit in itself, and is made up of a proportionate number of each of the nine digits in random order. The six series of 100 additions may then be considered as homogeneous and of approximately equal difficulty.

Subjects are not apt to comprehend this test so readily as they do, for example, the number-checking test. They may tend to add the numbers continuously, or more especially to add discrete numbers, thus with the column beginning as before, 8, 3, 5, 7, to announce the sums 11 and 12, omitting to add the 3 and 5. The operator must be well assured from the instructional material that the test is properly comprehended. Some subjects prefer to follow the columns with the finger; this should be permitted, on the same principle that permits reversed directions in the number-

3 3 1 9	4 3 2 4	5 6 4 5	5 5 3 2	5 4 6 5	9 5 6 5
6 2 5 9	5 6 6 5	2 9 5 3	4 2 6 9	3 8 6 6	3 7 7 8
5 5 8 6	8 9 9 2	6 3 6 7	3 1 2 4	4 6 2 3	5 9 3 2
1 2 2 7	9 1 3 1	7 1 1 8	2 5 6 3	7 2 1 1	1 3 4 5
3 5 3 9	5 7 7 3	4 7 9 5	5 1 9 6	8 9 4 3	4 6 4 7
3 8 6 1	4 5 6 8	5 3 6 3	1 8 1 1	1 3 4 5	3 8 8 3
2 3 8 5	2 1 2 7	5 7 7 3	4 9 4 8	2 5 7 2	5 1 8 7
7 7 7 9	7 5 7 7	2 2 4 8	5 2 7 8	3 3 6 2	8 6 9 6
6 2 5 5	4 5 1 1	8 4 9 4	2 1 3 5	3 8 7 6	2 1 5 5
5 8 5 4	2 2 1 9	9 6 2 2	2 7 9 4	8 9 5 4	4 2 3 2
8 4 3 8	5 5 8 9	5 7 3 9	6 8 7 9	5 5 9 7	8 1 5 6
9 4 1 2	9 1 6 7	1 7 7 1	7 6 9 7	1 3 7 1	6 9 9 7
9 1 1 9	8 8 5 3	2 5 5 2	9 8 1 8	5 6 8 4	3 8 7 4
8 2 5 7	1 3 8 4	7 5 6 3	8 8 5 3	2 4 1 5	9 7 2 4
6 5 9 8	6 5 2 8	2 3 8 4	6 7 7 2	3 1 1 4	9 3 5 7
4 5 4 7	6 4 1 4	1 8 1 6	1 1 9 9	9 1 5 2	6 1 4 7
3 8 4 1	3 6 9 5	7 1 8 4	2 3 6 6	9 6 8 5	1 9 8 1
7 5 5 2	2 4 6 8	9 8 9 3	4 4 5 7	6 7 5 7	6 5 2 3
2 5 1 4	1 8 9 1	3 5 1 8	7 6 2 5	5 9 7 8	2 9 2 1
7 4 9 3	7 2 7 4	2 9 4 5	5 8 5 9	8 3 9 6	3 4 5 6
1 6 8 3	8 6 2 4	6 2 1 1	4 5 3 7	4 2 1 9	5 4 8 1
5 1 6 7	5 3 5 7	1 5 5 6	5 3 4 5	1 8 4 9	5 1 9 7
2 6 9 7	6 8 2 7	9 8 7 6	3 5 1 6	5 9 9 2	2 7 2 4
4 7 6 4	3 9 9 3	4 3 8 5	3 2 8 1	6 7 5 4	9 8 5 8
6 9 1 6	6 5 3 4	9 9 6 4	9 7 8 5	8 3 8 7	1 6 4 5
2 8 3 4	5 3 5 9	5 4 8 2	4 3 4 6	2 2 7 5	3 5 2 6

checking test. The addition of the final digit of one column to the beginning one of the next is not required, nor do subjects ordinarily tend to do it, even though not specially instructed.

The test may of course be made of any length within the limits of the blank, but, from experience with the *Rechenhefte* one hundred additions appear sufficient for a unitary determination. The length of the present blank is believed to provide amply for sufficient variation of the experimental material for all ordinary purposes outside of special research upon the individual function.

The test requires the constant attention of the operator to the key to check the proper performance of the work by the subject. False reactions are more frequent than in the number-checking test, and they should be kept track of, though the writer has not seen an instance where they obscured individual differences in efficiency. The subject may not notice the error; if he does, the purpose should be to get him over it with as little disturbance and distraction as possible, and he should therefore be allowed to correct it or not, whatever is the path of least resistance for him. We believe this to be the sounder experimental practise, whatever might be said of it from an ethical standpoint.

2. In order to furnish a regulated experimental material which should have a greater flexibility of application than is usual in this class of tests, a second form of procedure is submitted, known as the *constant increment* test. This is a little-recognized method, but one which in direct comparison has shown superiority over other forms. It consists in presenting to the subject a series of numbers, requiring the identical arithmetical operation to be performed upon each. In the observations made with this test, the usual procedure has been the addition of 4. In this particular instance, there is perhaps no reason why the Kraepelin blank should not serve this purpose as well as that for which it is ordinarily used. In order however, to make the material adaptable also to subtraction, especially of larger figures, it is thought wise to preserve the special blank originally adapted to this test.

This blank contains 100 two place numbers. The unit places in these numbers are ten each of the figures 1,2,3,4,5,6,7,8,9,0. In

64	72	47	30
49	35	43	56
62	51	35	44
57	30	64	31
68	56	49	37
74	44	67	60
53	36	28	71
67	73	46	48
25	63	55	53
40	47	65	61
61	43	70	36
71	66	41	42
33	69	62	34
38	37	25	39
28	39	40	33
65	32	57	73
41	59	26	38
50	31	68	63
42	60	66	58
58	48	27	32
52	54	51	59
70	46	69	52
26	55	29	45
34	27	74	72
45	29	50	54

the tens places are ten 2's, ten 7's and twenty each of the intervening 3,4,5 and 6. These features are symmetrically distributed with reference to the halves. The subject speaks the proper responses according to the assigned experimental task; thus in the addition of 4 the responses begin 68, 53, 66, etc. It is probably advisable to confine the unit of observation to half of the blank, and even better would probably be two tests of one column each. Errors seem to be but slightly more frequent than in the simple addition of a pair of numbers; in their treatment, the same considerations obtain as in the previous form of test. Since the operation to be performed with the given numerals may vary indefinitely, no key is provided, but the experimenter may readily provide one himself for his own particular requirements, and should always do so; its employment being the same as in the Kraepelin form of addition test.

Results. Individual differences, due in part, no doubt, to differences of training, are very great in even the simplest arithmetical tests. Thus, while one of the authors has usually obtained, with the Kraepelin form of test, times of from two to three minutes for 100 additions, the other of us, working on 7 college and university students, has the following results:

Average time for 100 additions	107.2 seconds
A. D.	24.4 seconds
Range	65–164 seconds

With the constant increment test, the following results have been obtained from 10 subjects of the same class as above: Only one column was used in each test, and the times given are times for one column.

Problem	Add 4	Subtract 4	Add 17
Av.	33.9	41.1	97.4
A. D.	5.8	11.0	23.6
Range	24–49	25–67	62–158
Av. errors per column of 25	0.3	0.2	2.4

Experience with other subjects leads one to suspect that the time per column, for adding 4, will often run up above 60 seconds.

V. NAMING TESTS

1. *The Color Naming Test.* No doubt the best-trained and quickest of all associative responses appear in the reading of words, letters and numbers. (See above p. 9). Next to this, probably, stands the naming of familiar objects. The color-naming test belongs here, and may be expected to give shorter times, per reaction, than any of the other tests included in the present selection.

The test blank shows 100 patches of color (besides 5 sample patches), each 1 cm. square and separated by spaces of 1 cm. from its neighbors. The colors are to be named in order, as in reading, and one side is indicated as the top. The arrangement is such, however, that the blank can be equally well used in any one of the four positions.

Care is needed in the selection of colors for such a test; for the color names required should be of universal familiarity, and there are few such names. White, black, gray, red, yellow, green, blue, brown, and possibly pink and purple are in sufficiently general use for our purpose. It seemed safer to exclude pink and purple; if white is then used for the background, seven colors remain; but it seemed better, on the whole, to employ only five stimulus colors, repeating each twenty times. After some experimenting, it was decided to use red, yellow, green, blue and black.

The use of the Hering red and green seemed undesirable, since they do not suggest the color names so promptly as do a red and green each somewhat nearer to yellow. To attempt to equate the colors in brightness would not be best, since this would mean toning down the yellow to a point at which it would be likely not to be called by that name. The blue must be pale in order to be distinguished from black as quickly as the other colors are distinguished from each other.

In the arrangement of the color stimuli on the blank, all sequences of the same color were avoided, as such direct sequences

enable the subject to combine two reactions, practically, into one. Too frequent recurrence of the same sequence of different colors was also avoided, the object being to compel a separate reaction to each single stimulus. It was desired, indeed, to have the different possible sequences of two and of three colors occur with equal frequency; but something less than perfection in this regard had to be accepted, for it was desired also to make the arrangement equally good for each of four positions of the blank; and all the conditions of perfection could not be met in both the horizontal and vertical lines at the same time. An incredible amount of time was consumed in arranging the colors to meet these simple requirements. The blank seems now to be free from serious blemishes of arrangement, and requires approximately the same time for reading in each of the four positions.

Preparatory to the test, the experimenter lays the blank before the subject, with only the sample line of 5 colors showing. The subject is directed to give the names of the sample colors; when he understands the task and knows what names to use, the whole blank is exposed at the word "Go!" The time for the first half, as well as for the whole blank, may well be taken. Repeated trials can be made with the blank in four positions. The line of 5 sample colors is to be omitted in the actual test. As usual, one half of the blank is long enough for a satisfactory test. A key of the series of correct responses will facilitate the experimenter's task.

2. *The Form Naming Test.* The blank next to be described under the head of the Substitution Test is similar in all respects to the color sheet, except that five geometrical figures take the place of the five colors. This blank can be used in the same manner as the color blank, the numbers written in the key line being disregarded.

Results. Here, as in most of the other tests, the results now available are insufficient to do more than give a general impression of the time required. In this case, too, the results are all from one class of subjects, namely college and graduate students. The whole blank of 100 stimuli was reacted to, and the time taken for the half as well as for the whole.

Color naming test.

		1st half	2nd half	Whole
9 men	Av.	30.6	35.1	64.6
	A. D.	3.0	5.2	7.4
	P. E.	0.8	1.5	2.0
5 women	Av.	26.4	29.0	55.4
	A. D.	4.5	3.6	8.1
	P. E.	1.7	1.3	3.0
Total range, both sexes together		22-41	24-48	48-89

Form naming test.

		1st half	2nd half	Whole
6 men	Av.	46.7	47.2	93.8
	A. D.	8.6	7.2	15.5
	P. E.	2.9	2.4	5.2
4 women	Av.	38.5	42.3	80.8
	A. D.	6.0	10.0	16.0
	P. E.	2.5	4.2	6.7
Total range, both sexes together		31-60	29-58	60-117

From these data, it may be inferred (1) with reasonable assurance, that the color naming test is easier than the form-naming test. Comparison of the figures for the present color-naming test with those reported by Wissler[1] from the same class of subjects with the use of the Columbia color-naming test, which employs ten colors, makes it probable that the present test with five colors is noticeably easier—as was, indeed, intended.

(2) It may be inferred from the above table, with much probability, that a sex difference exists in the case of the color naming test, women being on the average quicker than men. This is the more probable because Wissler[2] obtained the same sex difference from much more extensive data.

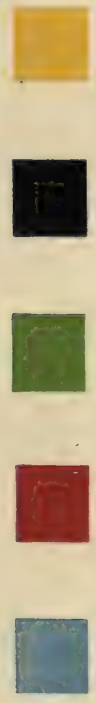
(3) It seems also probable that the same sex difference exists in the case of the form-naming test. If so, the sex difference here in question is not specially related to the color sense, but rather to linguistic facility. The authors have in mind the accumulation of sufficient data to determine whether the appearance here shown corresponds to a real sex difference.

(4) In the color-naming test it seems probable, and in the

[1] *Psychol. Rev.*, Monograph Suppl. No. 16, 1901.
[2] *Op. cit.*

form-naming test, possible, that there is some slackening of reactions, such as was shown above (p. 15) to occur in other forms of test. On reference to the individual records, we find that 10 of the 14 subjects in the color-naming test took longer for the second half than for the first half of the test; and in the form naming test, 6 out of 10 did the same. Only one of the exceptions is more than a bare exception. The behavior of the subjects during the test shows periods of hesitation and obstruction, and even of false reaction— rather a strange phenomenon, in view of the great familiarity of the names and their correct use immediately before. The subject is aware of this inhibition, and it is a strange experience for him. The "mechanism" of inhibition can not here have the elaborate Freudian character; and in fact the experiment seems a good one to show the reality of other forms of inhibition in the recall of names. The real mechanism here may very well be the mutual interference of the five names, all of which, from immediately preceding use, are "on the tip of the tongue", all equally ready and therefore likely to get in one another's way. These periods of inhibition do not appear at the very beginning of the test, but most often, to judge from incidental observation, along in the middle. Some subjects, after succumbing a few times to interference, appear to collect themselves and do the last part of the test better than the middle. We have the records of the successive rows of ten stimuli each, in the case of five subjects in each test. The average time in seconds for each row is as follows.

Row	1	2	3	4	5	6	7	8	9	10
Color naming ...	6.2	5.6	5.6	6.6	6.8	7.3	7.7	6.9	6.3	6.2
Form naming ...	8.0	9.1	9.0	10.3	10.8	8.2	10.5	9.9	10.3	9.0



VI. FORMATION OF NEW ASSOCIATIONS

The Substitution Test. This blank is modelled after one by Professor J. E. Lough,[1] but is simplified, in that only five (instead of twenty) different stimuli are used; at the same time, by employing geometrical forms in place of the letters of Professor Lough, it is partly freed from the danger that some subjects may hit upon easy mnemonics.

Since the names of the forms may enter into the subject's procedure, the forms should have equally familiar names. They should also be of such shapes that the blank, like the color-naming blank, may be capable of use in different positions. Only about five geometrical forms meet these conditions: the circle, square, triangle, star and cross. The blank is made up of these five forms, each repeated twenty times. The arrangement of the stimuli follows the same rule here as in the color-naming test.

At the top of the blank appears a line containing each of the five forms once, with a number on each. This line being covered, the rest of the blank is exposed to the subject, and it is explained to him that he is to write on each of these forms a number,— the same number as he will find on that form in the key at the top. In this test, the general rules of instructing the subject by aid of examples can not be exactly followed; for the association to be employed in the test should not be formed before the outset of the test itself, since this is a test of the formation of associations. When the experimenter is sure that the subject understands what is to be done, he uncovers the key, at the word "Go!" Besides taking the time for the whole and half, the experimenter may be able to get the times for each successive line, and so obtain a curve of the formation of the associations.

A misunderstanding which has occasionally appeared in the use of this test should be guarded against in the instructions to the subject. Some subjects have started to go through the blank numbering only one of the forms at a time, intending to

[1] Described by Kirkpatrick, Studies in Development and Learning, *Arch. of Psychol.*, 1909, No. 12, p. 36.

go through again for each of the other forms. It should be made clear that the forms are to be numbered in order as in reading or writing.

Results. Eleven educated adults (6 men, 5 women) gave the following average time in seconds:

	1st half	2nd half	Whole
Av.	79.6	65.1	144.7
A. D.	9.0	7.8	12.5
P. E.	2.3	2.0	3.1
Range	58-94	53-83	111-177

The gain from the first half to the second is perhaps not so great as would have been expected. In fact, few if any of the subjects fully mastered the key in the course of the 100 reactions.

Time was taken, in this test, for each successive row of ten forms, with the following average results (11 subjects):

Row	1	2	3	4	5	6	7	8	9	10
Av.	15.1	17.7	16.4	15.6	14.8	13.8	13.0	13.3	12.9	12.1
A. D.	3.7	2.8	2.8	1.6	1.6	2.3	2.2	1.9	2.4	1.9

The longer time for the first line than for the second is found in 9 of the 11 individuals; and 3 subjects do the first line as quickly as any other. It is, in short, possible to do the first line, by simple use of the key, in 9-11 seconds, and this is as rapid as any of the subjects became during the course of the test. Some subjects do the first line by mere copying from the key; others start to memorize and take longer on this line; this is probably the cause of the extra large variability for the first line.

The test is not long enough to permit the complete establishment of the associations; several blanks may be used in succession, and oral may be substituted for written responses in order to simplify the motor part of the performance. At the best, however, progress is rather slow; and, indeed, one would not expect these freshly formed associations to surpass readily the familiar associations involved in the form-naming test (p. 51), the times in which are, after all, not very much shorter than those in the last rows of the substitution test.

By use of a second key, the blank can be used for the study of interference; one of the authors has so used it in laboratory classes, following the general arrangement of Bergström's card-sorting experiment.[1] The blank can also be used for a simple cancellation test, similar to the number-checking tests.

[1] *Amer. Journ. of Psychol.*, 1893, *5*, 356.

VII. LOGICAL RELATIONS

The form of test in which the stimulus is a word and the response another word standing in some assigned logical relation to the stimulus has been long and widely used, and has an intellectual atmosphere that makes it seem likely to prove a test of individual differences of the intellectual sort. At the same time, it is distinctly a test of the command of language, and when the measurement concerns the speed of the response, familiarity with the necessary words is a prime necessity. Unless great care is used in the selection of stimulus words, long reaction times will occur from the need of searching for the proper response words, and the test thus becomes predominantly linguistic in nature. Linguistic it must always remain to a considerable extent, no matter how much care is taken in the selection of the stimuli; but the effort should be to minimize the linguistic factor by selecting only stimulus words that are universally familiar.

Besides the familiarity of the associations employed, the test calls for skill in the handling of these associations; and it is this skill, most of all, which the test designs to measure. In other words, it is the efficiency of the "determining tendency", or adjustment to react according to instructions, which should be revealed by the speed of performance. The more completely this adjustment dominates the performance, facilitating the right responses and inhibiting other, interfering associations and perseverations, the less hesitation and confusion will occur and the more prompt will be the reaction.

In order to afford sufficient opportunity for the determining tendency or adjustment to show its efficiency, it is customary and evidently desirable to provide a number of stimuli in succession, requiring the same sort of response to each. There should therefore be a list of stimulus words for each of the logical relations along which the reactions are to be required Thus the task of providing material for these tests consists in discovering a sufficient number of stimulus words of the requisite familiarity.

Only by actual trial can the suitability of the stimuli be ascertained. A word of apparently eminent fitness may prove to be unfamiliar to many subjects. For example, the word "false" seemed likely to be a good stimulus when the required response was a word of opposite meaning; but in practise much hesitation and uncertainty of reaction appeared in the responses to this word (instead of "true", some women subjects said "natural"). A word which seems perfectly familiar to the investigator may be unfamiliar to many subjects, and a word which seems perfectly unambiguous may convey an unexpected meaning to some subjects. To avoid all such difficulties with all subjects is too much to hope; but the test material should be freed from words that cause difficulty to a large share of the subjects. This is necessary at any rate if the time is to be taken, not for each separate response, but only for the whole series of responses to the list of stimuli; for otherwise the total time may be determined mostly by the difficulty of one or two of the reactions. Even if the times of the separate reactions are taken, the average time will suffer in the same way as the total time in the preceding case. The median time is mostly free from this source of error. But even so, lists of nearly uniform difficulty would form the best and fairest test material.

Our procedure in selecting stimuli for this class of tests was to start by getting together as large a number of stimuli as possible; to eliminate at once all that seemed ambiguous or unduly difficult and to try the remainder with a few subjects, timing the separate reactions, and eliminating the stimuli that gave the slowest reactions or that proved to be ambiguous or complex-arousing. The abbreviated list was tried in the same way with other subjects, and more words eliminated, till finally it appeared that the easiest possible list of stimuli had been secured. Unless it proved possible to secure a list of twenty easy stimuli, that particular test was abandoned. Thus, it seemed impossible to prepare a list of twenty words sufficiently easy for a synonyms test, except indeed for well-educated individuals. On the other hand, it appeared possible to select two lists of twenty for the opposites test.

The number of subjects employed in reaching the selection of stimulus words was greater in case of the opposites test than in the others, and this test may properly be regarded as more highly "standardized" than any other belonging under the head of logical relations. In the case of opposites, a long list of words was tried with 6 subjects, and the forty words selected from this trial were tried with 40 other subjects; some need of revision was then apparent, and a few more words were substituted from tests of a few subjects; then the revised collection of 40 was tried with thirteen fresh subjects, and a few minor corrections still introduced, which left the lists in their present condition. In the other tests, two to three times the desired number of stimulus words were tried with nine subjects, and the resulting selection was tried with thirteen fresh subjects; some minor changes were then introduced and the lists left in their present condition. The "mixed relations" test was selected gradually on the basis of results from fourteen subjects.

After the selection of the stimulus words came the question of their arrangement within the list. This matter of order of stimuli is not of great importance if the time is to be taken for the separate responses; but whenever the time is taken only for the series, the order of stimuli is a matter of some consequence. We recommend, it may be remembered, that the time be taken for the first half of the list as well as for the whole list, and even that the halves be given as separate tests; it is therefore important to have halves of equal difficulty. Moreover, many investigators find it convenient to allow a fixed time for each test, and to measure the number of responses that can be given in this time ("time limit method"); with this procedure it is important that the list shall be of uniform difficulty throughout, so that the number completed shall be a fair measure of the work done. Whatever be the procedure in giving the test, the most desirable arrangement of the stimulus words would be such as to distribute the difficulties evenly throughout the list. If it were really possible to discover twenty stimulus words of equal difficulty, the question of their arrangement would not arise; but this is not possible, for though the twenty stimuli be all decidedly easy, yet the reaction time to one will be two or three times as long as to

another of the twenty, on the average of as many as ten subjects. Since it is impossible to prepare a list of twenty stimuli of equal difficulty, we combined the words in pairs so that the pairs should be of equal difficulty, as judged by the sum of the reaction times to the two members of each pair. One pair thus may consist of the hardest and the easiest word in the list, and another pair of two words of medium difficulty; but the sum of the reaction times for the first pair is equal to that for the second pair, as judged from the records already in hand. (The pairs can not be hoped to be equal for all subjects.) These equal pairs can then be arranged in such a way as to avoid "constellations" or undesirable collocations of any sort; and the difficulties of the list will be pretty evenly distributed.

Two other points were considered in arranging the order within the lists. When the test is given with a time limit, it is especially desirable to have the responses of uniform difficulty in that part of the list where most of the subjects will be stopped, so that there, at least, the single words shall constitute equal units. We have therefore placed those of our "pairs" which are composed of words of medium difficulty in the midst of the list, from about the 8th to about the 16th word. If then the time limit is so chosen that the great majority of subjects shall be stopped in this part of the list, the separate words may, without much error on the average, be counted as equal units.

The other point concerns the writing of responses. In reality, as explained in the introduction, an easy association test is very ill adapted for written responses, because the time of writing is much greater than that of easy association, and individual differences in speed of writing altogether mask the differences in speed of association. However, in case of the opposites test, we have determined the writing times for the correct response and so distributed the stimuli that the writing times for the two lists of twenty, for the halves of each list, and, as nearly as possible, for the pairs throughout each half, shall be equal. This has not been attempted for the other tests given below, because the response words are not wholly determined in advance.

1. *The Opposites test.* This test has one advantage over all the others in the series of logical-relation tests, namely that the

answers can be definitely scored as right or wrong. Opposites apparently are the most available material for a test of completely controlled association—with the exception, indeed, of the naming tests and of the arithmetical tests already brought forward. For this reason, we have taken unusual pains with the selection and arrangement of material for this test. As it appears possible to select forty words free from difficulty, we offer two lists of twenty, with the object of making it possible to give two equally difficult tests of the same function. Since, however, it may be desired in some instances to have the very simplest material, we also present a list of twenty "Easiest Opposites", all of which are included in the two lists of twenty.

The lists are printed on three separate slips,[1] in 12-point type, well-leaded. Lists I. and II. are of equal difficulty, and the halves of each list are equivalent, as far as can be judged from the results of their use so far.

The instructions, enforced by samples (see p——), require the subject to respond to each stimulus word by the word having the opposite meaning; as, "long-short".

OPPOSITES TEST

I	*II*	*Easiest*
long	north	high
soft	sour	summer
white	out	out
far	weak	white
up	good	slow
smooth	after	yes
early	above	above
dead	sick	north
hot	slow	top
asleep	large	wet
lost	rich	good
wet	dark	rich
high	front	up
dirty	love	front
east	tall	long
day	open	hot
yes	summer	east
wrong	new	day
empty	come	big
top	male	love

[1] The type, etc., of these tests is not reproduced here.

2. *The verb-object test.* More good stimulus words are available here than in any similar test except that for opposites. Here again we have selected two equivalent lists, and also a list containing the very easiest stimuli, as judged from results with about 20 subjects. The verbs are to be treated as transitive, and objects supplied; for example, "sing song", "build house." The audible repetition by the subject of the stimulus word is not required, and may interfere somewhat with the experimenter's record; but it does not change the times to an appreciable extent.

VERB-OBJECT TEST

I	*II*	*Easiest*
sing	read	wash
build	tear	sing
wear	throw	bake
shoot	paint	read
scold	mail	chew
win	light	learn
answer	sail	mail
weave	spin	sweep
wink	lock	scold
mend	wash	wear
pump	bake	sharpen
learn	spill	kiss
open	kiss	smoke
eat	polish	answer
climb	sweep	climb
lend	fill	lock
smoke	sharpen	throw
singe	write	sail
dig	chew	dig
sift	drive	wink

3. *The supraordinate concept or species-genus test.* The instructions are to name a class to which the given object belongs, or to "tell what sort of thing each is"; as "oak—tree."

4. *The subordinate concept or genus-species test.* The instructions are to name an example of the class mentioned, or to "mention a—", as, "color—red".

5. *The part-whole test.* The instructions are to name the whole thing of which the part is mentioned; as "elbow—arm".

Supraordinate Concept Test	*Subordinate Concept Test*	*Part-Whole Test*
oak	color	elbow
measles	holiday	hinge
July	fish	page
shark	tool	finger
quinine	metal	wing
beef	vegetable	morning
canoe	coin	blade
banana	city	mattress
Atlantic	insect	chimney
Alps	food	cent
penny	fruit	sleeve
dictionary	disease	brick
cabbage	grain	deck
Rhine	drink	France
murder	month	pint
dog	ocean	fin
sparrow	language	steeple
London	river	month
football	newspaper	hub
rose	tree	chin

Whole-Part Test	*Agent-Action Test*	*Action-Agent Test*	*Attribute-Substance Test*
apple	baby	gallops	sharp
clock	fire	bites	hot
knife	dog	boils	dusty
book	laborer	sleeps	raw
hat	pencil	floats	deep
pencil	army	growls	ripe
hand	heart	sails	funny
dog	pin	roars	tall
oyster	gun	scratches	stormy
church	eyes	stings	new
chair	bird	shoots	hilly
bird	wind	melts	strong
banana	lungs	swims	muddy
shoe	bell	explodes	pretty
train	musician	aches	noisy
finger	parrot	blows	white
house	clock	mews	steep
coat	axe	cuts	round
cart	broom	flies	smoky
face	mosquito	burns	curly

6. *The whole-part test.* The instructions are to name a part of each thing mentioned; as "apple—core".

7. *The agent-action or subject-verb test.* The instructions are to put an appropriate verb to each noun as subject; or to "tell what each of these does or can do;" as "baby—cries".

8. *The action-agent or verb-subject test.* The instructions are to supply a subject to each verb, or to "tell what does or can do each of these things;" as, "horse gallops."

9. *The attribute-substance or adjective-noun test.* The instructions are to supply an appropriate noun for each adjective, or to "tell something that is or may be each of the following", or to complete the expression, "A good—", etc.; as, "sharp knife",

10. *The mixed relations test.* In the preceding tests, the task remains the same through a series of reactions; in the present test the particular relation along which the reaction is required to occur changes with each reaction—the object being to get some insight into flexibility of mental performance. We were long at a loss for some means of indicating the new task without lengthy explanations at each new stimulus and also without the use of such technical terms as supraordinate, etc. Finally a device used by one of us previously in the study of consciousness of relations seemed to meet our present needs: the relation along which the reaction is to take place is indicated before each new stimulus word by a pair of words serving as a sample. The subject is to note the relation of the second word to the first, and then find a word standing in this same relation to the third word. Thus, in the example "Box—square Orange—?" "square" gives a quality of "box", or, more specifically, the shape of the box, and it is required to mention the shape of an orange; in the example "East—west Day—?" since east and west are opposites the task is to find the opposite of day; and in the example, "Penny —copper Nail—?" the task is to mention the material of which the nail is composed. Some of the relations are not readily named, but little difficulty has appeared, with the adult educated subjects already tested, in grasping the relation from the sample given. Instructions for this test must proceed largely by

the use of samples, of which several must be given, in order that the subject may realize that it is not always the same relation that is needed, but a new relation each time as indicated by the first two words in the line.

Mixed Relations Test
I

Eye—see	Ear—
Monday—Tuesday	April—
Do—did	See—
Bird—sings	Dog—
Hour—minute	Minute—
Straw—hat	Leather—
Cloud—rain	Sun—
Hammer—tool	Dictionary—
Uncle—aunt	Brother—
Dog—puppy	Cat—
Little—less	Much—
Wash—face	Sweep—
House—room	Book—
Sky—blue	Grass—
Swim—water	Fly—
Once—one	Twice—
Cat—fur	Bird—
Pan—tin	Table—
Buy—sell	Come—
Oyster—shell	Banana—

Mixed Relations Test
II

Good—bad	Long—
Eagle—bird	Shark—
Eat—bread	Drink—
Fruit—orange	Vegetable—
Sit—chair	Sleep—
Double—two	Triple—
England—London	France—
Chew—teeth	Smell—
Pen—write	Knife—
Water—wet	Fire—
He—him	She—
Boat—water	Train—
Crawl—snake	Swim—
Horse—colt	Cow—
Nose—face	Toe—
Bad—worse	Good—
Hungry—food	Thirsty—
Hat—head	Glove—
Ship—captain	Army—
Man—woman	Boy—

Results with the logical relations test. After the tests had reached practically their present condition, they were tried with thirteen college and graduate students (in a few cases, the number of individuals was less than this). In these experiments, lists of ten stimulus were presented visually, but the time of the single reactions was roughly taken by the device mentioned on page 17. Usually two, and in the case of the opposites and verb-object tests four lists of ten were used, and each subject's average time per single reaction was obtained. The averages given in the accompanying table are the average of the individual averages, and the A.D. is that of the individual averages from the general average.

	Av. per single reaction	P.E.	A. D. of indivs. from general av.	Range of indivs.
Opposites 1 and 2.....	1.23	.06	.16	1.03—1.50
Opposites, easiest	1.11	.04	.12	0.85—1.40
Verb-obj. 1 and 2.....	1.39	.05	.19	1.08—1.75
Verb-obj., easiest	1.31	.05	.14	1.10—1.55
Suproord. concept.....	1.54	.07	.31	0.90—2.20
Subord. concept.....	1.84	.07	.31	1.20—2.63
Part-whole	1.53	.06	.27	1.03—2.50
Whole-part	1.57	.07	.32	1.13—2.35
Agent-action	1.30	.03	.12	0.93—1.70
Action-agent	1.55	.07	.32	1.03—2.68
Attrib.-subst.	1.53	.07	.28	1.08—3.05
Mixed relations	3.14	.13	.53	2.33—4.40

The degree of agreement between the results of the several logical relations tests is a matter of some interest as indicating to what extent a single test is a fair indication of the individual's ability in this whole class of performances. By methods which will be more fully described in another paper, we have determined the average standing of each of our thirteen subjects in the nine logical relations tests (excluding the mixed relations test), and have correlated this average standing with the standing in each single test. The results follow, in the form of Pearson coefficients, uncorrected for attenuation.

		r	P.E.
Correlation of Average with:	Opposites.....	+.88	.03
	Verb-object.....	+.70	.08
	Subordinate conc.	+.72	.07
	Supraordinate concept....	+.91	.03
	Part-whole	+.86	.04
	Whole-part	+.76	.06
	Agent-action	+.83	.04
	Action-agent	+.84	.04
	Attribute-substance	+.54	.12

As far as these few results indicate, then, the opposites and supraordinate concept tests seem slightly better than the rest as representative of this general sort of controlled association. The correlation between the opposites and the supraordinate concept tests was +.70, with P.E. of .08, while the average correlation between any two of the nine logical relations tests is +.57.

Comparative speed of the different forms of controlled association

It may be of interest, since results are available in several tests from comparable and highly reliable (though not numerous) subjects, to bring together the *times per single reaction,* placing them in the order from quickest to slowest. In all these cases, a series of stimuli was simultaneously presented, so that overlapping took place.

Performance.	Time in seconds per single reaction.	P.E.
Reading letters or short words	0.33	
Naming colors	0.61	.03
Naming forms	0.89	.09
Adding two one-place numbers (Kraepelin blank)	1.07	.08
Easiest opposites	1.11	.04
Agent-action	1.30	.03
Verb-object, easiest	1.31	.05
Adding 4 (const. increment test)	1.36	.08
Part-whole	1.53	.06
Attribute-substance	1.53	.06
Supraordinate concept	1.54	.07
Action-agent	1.55	.07
Whole-part	1.57	.07
Subtracting 4 (const. increment test)	1.53	.07
Subordinate concept	1.84	.07
Mixed relations	3.14	.13
Adding 17 (const. increment test)	3.90	.31

These differences are of course not to be understood as meaning that the finding of opposites is always, or even on the average, quicker than the finding of wholes when parts are given. The times for opposites that are by no means recondite or unusual run up to an average of at least 5 seconds per reaction. It would be futile to attempt to determine the average or median time for all opposites, and even more futile to make such an attempt in case of the part-whole, genus-species and many other relations; there would be no way of setting an upper limit to the difficulty of the single stimulus words. Such a statement as that the mind passes more readily from species to genus than from genus to species has therefore not much real validity. The fact simply is,

as far as our results are concerned, that the easiest opposites are easier than the easiest part-whole associations, etc.; and by "easiest" is meant, in case of the several logical relations tests, the twenty easiest. The differences between the speed of controlled associations are perhaps mainly dependent on the factor of frequency in past experience, and especially on the frequency of linguistic transitions. Thus, transitions between opposites are frequent in common speech, and many pairs of opposites thus become verbally associated in a high degree. The reproductive tendencies in case of the most commonplace opposites are therefore strong; and it may also be that the "mental set", or "determining tendency", is better drilled in case of finding opposites than in many other sorts of logical relation.

VIII. THE UNDERSTANDING OF INSTRUCTIONS

As already mentioned (p. 20), a test should not ordinarily be begun till the subject certainly understands the instructions; otherwise the time measured is partly occupied with grasping the problem, and only partly with its execution. Each test should, as nearly as possible, be a test of one sort of performance. But it seemed desirable to attempt to test the ability to understand instructions, and accordingly efforts were made to prepare a test which should give many different sets of very simple instructions, with the object of discovering the subject's speed in apprehending them. After much experimenting, the following were produced, and the test was named the *directions test.*. This test should, we believe, be given as a list or continuous test, with rough timing also of the single reactions, so as to get the median as well as the average time of response. The reactions are to be made with a pencil; and the test can very well be made with a time limit as well as with an amount limit.

The conditions which it was sought to meet in the test material are (1) that the motor response should be very simple and quickly performed; (2) that the instructions should be very simple, but varied; and (3) that the instructions should be as concise as possible, in order that reading time might not be the determining factor.

1. *Easy directions test.* Two blanks are provided, of approximately equal difficulty, according to the results so far in hand. The halves are also approximately equivalent.

2. *Hard directions test.* The object here is to complicate the directions somewhat, by calling for conditional and alternative responses, etc. The blank is arranged in the general form of an Ebbinghaus combination test. The instructions are simply to fill in the blank according to the directions in it. The first two or three directions are easy, so as to put the subject on the right track. The remaining units within the blank (except the last) are so chosen as not to be very unequal, with the object of making

the blank available for use with a time limit. It can not, however, be claimed for this test that it is as well worked over and standardized as the others in this series.

Results with the directions tests. Data so far in hand are rather meager, eight subjects having taken the easy directions test in approximately its present form, and six subjects the harder test—all educated adults. The results follow:

DIRECTIONS TESTS.

	Av.	P.E.	A. D.	Range
Easy tests, time in secs. per reaction..	3.60	.28	.92	2.30—5.70
Hard test, time for whole blank 107.6		6.0	18.4	76—134

If the number of reactions in the hard test is counted as 20 (which is approximately correct), the average time per reactions is 5.38 seconds; the reactions are no doubt slower in this than in any other of the tests described in this paper. To judge from the six subjects who have taken both the easy and the hard directions test, the correlation between the two is very high (Pearson $r = +.92$).

Cross out the smallest dot: • • •

Put a comma between these two letters: G H

How many ears has a cat?

Make a line across this line: |

Show by a cross which costs more: a hat or an orange.

Write 8 at the thinnest part of this line: ——

Write any word of three letters.

Put a dot in one of the white squares: 

Cross out the word you know best: fish, brol, matzig.

Leave this just as it is:  

Mark the line that looks most like a hill: 

How many t's are there in twist?

Dot the line that has no dot over it: 

Write o after the largest number: 3 86 72

Mark the name of a large city: London, painter.

Make a letter Z out of this: 

Join these two lines: —— ——

Write s in the middle square: 

Write any number smaller than 10.

Put a question mark after this sentence

Cross out the *g* in *tiger*.

Write 2 between the two dots: • • ——

How many feet make a yard?

Write + over the longest word. It rained yesterday.

Put a dot below this line: ——————

Write the sum of these numbers: $\frac{3}{4}$

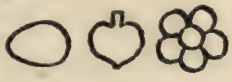
Make a boy's name by adding one letter to Joh

Make a cross in the circle: 

What comes next after D in the alphabet?

Write 7 in the largest square: 

Cross out the blackest letter in TEXAS

Write *g* on the egg-shaped figure: 

Make two dots between these lines: 

Put the sign = where it belongs: 3 + 2 5.

Write here..... the middle letter of *get*.

Put a nose on this face: 

Add a cross and make these rows equal: X X X
X X X X

Put a dot in the circle, below the center: 

Draw a line around the three dots: • • • • • • • •

Cross out the last word in this sentence.

With your pencil make a dot over any one of these letters **F G H I J**, and a comma after the longest of these three words: **boy mother girl** Then, if Christmas comes in March, make a cross right here..... but if not, pass along to the next question, and tell where the sun rises..... If you believe that Edison discovered America, cross out what you just wrote, but if it was some one else, put in a number to complete this sentence: "A horse has.....feet." Write *yes*, no matter whether China is in Africa or not .....; and then give a wrong answer to this question: "How many days are there in the week?"..... Write any letter except *g* just after this comma, and then write *no* if 2 times 5 are 10..... Now, if Tuesday comes after Monday, make two crosses here.....; but if not, make a circle here.....or else a square here .... Be sure to make three crosses between these two names of boys: George.....Henry. Notice these two numbers: **3, 5.** If iron is heavier than water, write the larger number here....., but if iron is lighter write the smaller number here..... Show by a cross when the nights are longer: in summer?..... in winter?..... Give the correct answer to this question: "Does water run uphill?"..... and repeat your answer here..... Do nothing here ($5 + 7 =$ .....), unless you skipped the preceding question; but write the first letter of your first name and the last letter of your last name at the ends of this line:

IX. THE FREE ASSOCIATION EXPERIMENT

Few procedures in experimental psychology have so richly rewarded their investigators with the possibilities of practical application. In ordinary psychological nomenclature, it is *the* "association" experiment *par excellence*. Within the past seven years it has achieved, and bids fair to hold indefinitely its place in the foremost rank among the methods of individual psychology. The body of work that has gathered about it is probably greater than that about any other single psychological experiment, and it is not surprising that it constitutes one of the best understood, as well as one of the most potentially significant of them.

The preliminary task of standardization is to provide as error-free a method as practicable, but the main object of standardization is to afford a basis for making comparisons between different individuals. An experimental method becomes standardized in the most complete sense when, given a proper technique, it is possible to accurately rate individual records with reference to an empirical scale. None of the "mental tests" possesses this quality to a degree comparable with the free association experiment, within the limits of the English language. This is mainly due to the work of Kent and Rosanoff which established a definite standard of normality for a specific association material.[1] Within the bounds of its application, it would be an impertinence to offer as "standard" any procedure for the free association test other than the one which these authors have developed; our first endeavor will be then to describe this experimental material, and to indicate what seem to be the best methods for its application.

The Kent-Rosanoff experiment consists of one hundred ordinary English words of somewhat varying difficulty, in the order given on the opposite page, and the making of the test

[1] Kent and Rosanoff, A Study of Association in Insanity, *Am. Journal of Insanity*, LXVII, pp. 37-96 and 317-390.

requires, according to the subject, from ten to twenty-five minutes.

This experiment was made by Kent and Rosanoff with 1000 normal subjects, and the responses were tabulated to each individual stimulus word. These constitute the so-called "Frequency Tables," and their use is to determine the "value," in terms of its frequency coefficient, for any reaction or series of reactions in a record of this experiment. After each response-word in the tables occurs a number, which is the number of times that the word to which it attaches occurred as a response to the stimulus word among the 1000 observations. This figure, divided by 10, is taken as the "value" of the response. Thus the "value" of the association *table-accommodation* is .1, because it was given by only one of the 1000 subjects; that of *table-chair* is 26.7, being given by 267 of the 1000 subjects, that of *dark-room* is 2.2, that of *music-art* is .7. It is found, then, that different records of the test show marked differences in the "value" or *usualness* of the associative responses. By means of these frequency tables, the proper "value" is assigned to all reactions obtained from the stimulus-words. ***Any one wishing to work with this experiment must provide himself with a copy of the tables,[1] which it is impossible to reproduce here.*** Some measure of central tendency for all the measures should be taken, and the distribution of the measures indicates the median to be preferable for this purpose to the average, aside from its greater ease of calculation.

The first and foremost datum of the Kent-Rosanoff experiment is an empirical measure of the tendency of the subject's train of thought to move in usual or individual channels; more accurately speaking, along objective or subjective lines. A number of interpretational questions arise in connection with this finding, which seems less correlated with education than with temperament. It is perhaps the best objective correlate of temperament at present to hand, but the matter is a rather complicated one, more suitable for separate discussion. Here need be emphasized only the preciseness and objectivity with

[1] *American Journal of Insanity*, LXVII, pp. 48-90. To be had of G. E. Stechert & Co., New York.

1. Table	26. Wish	51. Stem	76. Bitter
2. Dark	27. River	52. Lamp	77. Hammer
3. Music	28. White	53. Dream	78. Thirsty
4. Sickness	29. Beautiful	54. Yellow	79. City
5. Man	30. Window	55. Bread	80. Square
6. Deep	31. Rough	56. Justice	81. Butter
7. Soft	32. Citizen	57. Boy	82. Doctor
8. Eating	33. Foot	58. Light	83. Loud
9. Mountain	34. Spider	59. Health	84. Thief
10. House	35. Needle	60. Bible	85. Lion
11. Black	36. Red	61. Memory	86. Joy
12. Mutton	37. Sleep	62. Sheep	87. Bed
13. Comfort	38. Anger	63. Bath	88. Heavy
14. Hand	39. Carpet	64. Cottage	89. Tobacco
15. Short	40. Girl	65. Swift	90. Baby
16. Fruit	41. High	66. Blue	91. Moon
17. Butterfly	42. Working	67. Hungry	92. Scissors
18. Smooth	43. Sour	68. Priest	93. Quiet
19. Command	44. Earth	69. Ocean	94. Green
20. Chair	45. Trouble	70. Head	95. Salt
21. Sweet	46. Soldier	71. Stove	96. Street
22. Whistle	47. Cabbage	72. Long	97. King
23. Woman	48. Hard	73. Religion	98. Cheese
24. Cold	49. Eagle	74. Whiskey	99. Blossom
25. Slow	50. Stomach	75. Child	100. Afraid

which it is possible to evaluate an experiment of such intimate and subjective character.

If special circumstances render it desirable, it is possible to employ only a larger portion of the hundred words for determination of the usualness in response, substituting for the remainder, words adapted to the special situation in hand. It would be desirable indeed, if the Kent-Rosanoff experiment were made the framework of all experiments for *Tatbestandsdiagnostik*, the individually significant words being either added to it or replacing some, not over 10 or 15 per cent. of its constituents. To deal objectively with questions of *Tatbestandsdiagnostik* requires a number of precautions in the construction of the special series, the enumeration of which would be out of place here, and which are fully discussed by the investigators of this application of the method.

Unfortunately, determinations of the "median of community" (i. e., the median "value" of the 100 reactions in a record) have as yet been made in only a limited nnmber of subjects. In some pathological cases it would become indefinitely small; the lowest median ever observed by either of the writers in a normal subject is .7. Such a figure would mean that half the reactions of this record were of a frequency below that of the reaction *music-art* quoted above. The other extreme of the range, so far as observed, is 18.2, i. e., half of the reactions in such a record are more common than, i. e., *music-piano*. The general average value of the reactions in the above mentioned records lies not far from 9.0, that is, about the frequency of a reaction such as *mountain-valley*.

The present experimental method is placed under one disadvantage to a much greater degree than other association tests; its material cannot be repeated within an ordinarily practicable time save under greatly changed essential conditions. One can foresee that circumstances may arise in which a comparative study with material of greater extent is desirable. Provision is here made for such material to be available,[1] but with a change in the character of the material comes inevitably a change in the method of evaluation. Beyond the range of the frequency tables

[1] See Appendix, pp. 80 ff.

one must fall back on the quasi-logical system of classifying the associations that was practically the sole means of dealing with such material until the data on statistical frequency were compiled. The proper function of the test, however, is the same as before, and so is the object of its evaluation: the measure of egocentricity in the responses.

There is no need to fully repeat the remarks in a previous contribution regarding the method of evaluation that seems best adapted to these conditions. It is a five-fold classification, including categories termed (1) the egocentric, (2) the supraordinate, (3) the contrast, (4) the miscellaneous or "internal objective," and (5) the speech-habit.[1]

For ordinary purposes of comparison, the principal question concerns the number of reactions that fall into the category of the egocentric; and a large or small number of such associations is subject to analogous interpretations with the empirically determined tendency towards common or individualized responses.

[1] The definitions and illustrations of the categories may be summarized from the previous paper as follows:

1. The egocentric reactions may be typified by—

a. Predicate reactions. *Cloud-ominous, flower-pretty, crooked-line, red rose, scratch-cat, lion-roar, money-wish, invent-machine, weasel-stealth, beauty-rose, safe-quite, almost-grown, sing-well, never-decide, nicely-very* (including the responses yes and no).

b. Responses in the form of proper names. *Citizen-New York, boy-Johnny, mountain-Kearsarge.*

c. Reactions interpreting the stimulus word as a proper name. *Eagle-newspaper, park-square.*

d. Reaction involving the response of a pronoun. *Hand-you, health-me.*

e. Interjections, failures of response or repetitions of the stimulus word.

2. The supraordinate category is confined strictly to the individual-genus order, defined in such examples as, *priest-man, potato-vegetable, lily-flower, cow-animal.*

3. The contrast group is composed, of course, of reactions in which the response meets the opposite of the stimulus and is made up of such associations as, *good-bad, trouble-pleasure, scatter-gather, fertile-sterile,* and the like.

4. The miscellaneous category is composed essentially of the remaining reactions of the "inner" type. It includes about 45% of all associations.

5. The speech-habit group is composed of associations by familiar phrase (*stand-pat*), word compounding (*play-ground*), simple sound associations (*tease-sneeze*) and syntactic changes (*high-height*). (*Psychol. Review,* 1911, *18,* 229-288.)

The egocentric is the most variable of the five categories, ranging from next to none to more than half of the total reactions in a single experimental series.

With respect to timing the associations, the stopwatch is the almost universal method, and must be accepted as such, until some more accurate procedure is devised. Significant differences are usually coarse enough to be manifested in measures of no greater precision than this. More uniformity in the manipulation is desirable; at present, some operators start the watch on the accented syllable, others at the first syllable, of the simulus word. The watch should always be stopped at the first indication of response, even if it does cause occasional failure of timing through the subject's clearing his throat.[1]

Individual differences in association time should be discussed from the standpoint of the distribution rather than any single measure. The median is rather preferred as a measure of central tendency, though for practical purposes, its advantages over the average are of less account here than in most cases of skew distribution. The presence of many and exclusively long measures happens to be more important here than in most similar series of measures. Jung has proposed a special comparison of the average and median; this is a convenient statement of the distribution, but it is not an index of emotivity, beyond the limited extent to which the association time can be interpreted in this direction.

The usual instruction in the free association test is that the subject shall reply with the first *word* the stimulus suggests to him, but in ordinary practise this is not rigidly enforced, it being sometimes possible to derive elements of special significance from factors that determine the subject's departure from the set instructions. For comparison with the frequency tables, it is readily apparent that the single word response must be rigidly required in the Kent-Rosanoff experiment; in cases of dereliction from this rule, it is the practise of these authors to repeat

[1] Coughing at such times has received some notice as a *Komplexmerkmal*, though it has been sagely remarked that this loses much of its significance if the subject in question has a cold.

the stimulus word at the end of the experiment, in order to obtain a reaction of the required character.

Presentation and response have nearly always been oral in this experiment, and there is good reason to make no change in them. Experimenters differ as to the manner in which they modulate the voice to the test. Some experimenters, as Dr. Rosanoff, speak the stimulus word with a rising inflection, as though asking a direct question, some as an exclamation, as though endeavoring to hurl the subject's "complexes" in his teeth, others in a monotone. There are subjects whose responses might be somewhat affected by these differences in procedure. This is one of the minor features of the test, in which, though uniformity among different investigators is scarcely practicable, the experimenter should at least strive to keep his own procedure constant.

And while it is not proven, it is most probable that the responses are also affected by the personality of the operator. In making fine comparisons between records by different experimenters this fact must be borne continually in mind.

So far as developed, the continuous form of the free association experiment is a method of possibilities rather than of promise. It would probably be capable of performing many of the functions of the discrete form, but there are external reasons why it would not be likely to perform them so well. The problem of standardization would be practically confined to the development of the most significant scheme of evaluation after the pattern indicated by previous workers with the test. It might be possible to employ the scheme of evaluation proposed for the discrete free association experiment, each single word given being allowed to serve as the stimulus word for the next association. So far as normal psychology is concerned the method has thus far dealt very largely with group averages. As a method of individual psychology it may assume a position comparable to its better known congener only as a result of extended and laborious researches.

NOTE: Blanks for all the preceding tests are to be obtained from the C. H. Stoelting Company, 113-125 North Green St., Chicago.

APPENDIX

The following series of 1000 words is intended for general use in the free association experiment. It is a revision of the series employed in the experiments on the practise effects in the test, and is modified in the manner suggested by the experience of this investigation. It is intended to contain 1000 different words, none over three syllables, so far as possible familiar and unambiguous. It is not far from exhausting the total available number of such stimulus words. Ambiguous stimulus words have a special and useful purpose, but not in a test of the present character. The details of the preparation of this list were substantially the same as in the previous list, save in one particular. The division into twenty series of fifty words each is followed here. But the present list also contains the hundred words of the Kent-Rosanoff series, distributed pro rata, five words in each series, and in their actual order of sequence in the Kent-Rosanoff test; otherwise their arrangement in the series is random, save that none occur in the first ten words of a series. In the list as printed, the words from the Kent-Rosanoff series are distinguished from the remainder by an asterisk, and the associations of these words may be evaluated by the frequency tables. In using single series of fifty words it is recommended that a sheet of paper of fifty lines be obtained upon which several records of reactions to the same stimuli may be conveniently noted. Stimulus words not evoking a reaction according to instructions may be repeated at the close of the series; and if a stimulus word evokes, as a response, the word coming next in the series, this word is omitted and given at the close.

The complete list is as follows:

I	II	III	IV
bottle	drink	locust	weary
produce	captain	divide	tooth
rope	cedar	restore	practise
delicate	mischief	tempt	supper
thick	clean	fade	tun
end	also	cheap	pepper
omelet	path	compel	best
expensive	ride	power	heart
cap	salute	baker	island
barrel	grocery	athlete	machine
burglar	bashful	*black	pit
design	true	roof	*fruit
cry	perverse	cradle	return
hip	occasion	certain	marriage
overcoat	nuisance	travel	marsh
freeze	*deep	impress	owl
*table	pinch	daughter	water
lightning	satisfy	gun	summer
follow	tank	book	copper
parlor	hat	barber	beetle
smoke	nourish	natural	statue
stretch	sister	elephant	clothes
tar	*soft	ostrich	oblong
snake	ham	curse	*butterfly
purpose	ugly	*mutton	constable
*dark	age	haste	cloud
unfair	glory	lizard	collapse
ditch	tough	result	solid
tiger	acid	nonsense	number
*music	*eating	index	goose
wicked	crowd	fool	railroad
prefer	discourse	dense	excite
fish	watchful	life	hornet
instrument	indecent	wine	*smooth
guilty	exchange	*comfort	delay
seed	costume	fever	begin
*sickness	style	infirm	cat
crush	trap	comb	asylum
rich	*mountain	spice	knee
hash	drift	starch	tight
unseen	crime	venture	car
death	cover	*hand	*command
umbrella	abuse (v)	pirate	insect
blood	open	brandy	hope
gift	*house	dress	insist
*man	enjoy	pebble	*chair
allow	untrue	adventure	star
sailor	dismay	lip	ice
prospect	unburden	*short	picture
school	again	pint	bind

V	VI	VII	VIII
room	crown	forget	pot
pencil	get	goat	camp
dig	honest	pulse	shirt
indiscreet	vacant	unwholesome	chain
restless	beechnut	attention	ádult
simple	splinter	dairy	violin
measure	unbelief	boast	reason
loss	argue	color	excuse (n)
reckless	cónflict	chin	roast
flirt	alike	servant	fig
pause	skin	dislike	face
prosper	inside	dead	*red
*sweet	*wish	*rough	common
avoid	hero	fortune	complexion
fresh	scarlet	candy	deserve
real	lamb	perfect	dim
potato	neck	disdain	view
*whistle	clasp	fierce	*sleep
bite	spear	constant	dirt
clear	*river	violent	shoe
eternal	ox	care	slave
jealous	serious	indeed	protect
barn	garter	*citizen	sting
*woman	key	death	funny
persuade	conquest	sparrow	solemn
merit	*white	*foot	little
receive	scratch	over	*anger
above	cool	bother	ramble
conceal	correct	forward	family
revolt	paste	prepare	annoy
*cold	uncertain	establish	confusion
join	pudding	gold	ripe
tender	*beautiful	along	greasy
offense	bacon	cannon	admire
guide	rancid	boat	cup
prompt	fertile	song	insult
floor	dog	*spider	easy
advance	perfume	another	impudent
bundle	toy	irony	*carpet
ignorant	dust	art	decay
*slow	pansy	dove	*girl
blunder	lake	poem	announce
confidence	*window	herald	prudent
knob	vow	*needle	trumpet
flesh	pancake	treasure	convenient
future	cork	sensitive	ask
wart	gay	fog	supreme
market	feast	tunnel	portable
chocolate	gem	remain	before
disease	enormous	closet	alone

IX	X	XI	XII
accuse	crab	uphold	intimate
flame	cart	pickle	able
clown	riot	food	suspect
edge	preserve	raven	barley
frost	preach	unwell	attack
rust	unclean	ready	dishonor
corrupt	steep	away	accident
appear	master	blame	betray
caution	sulphur	competent	door
poverty	laugh	mark	prince
curtain	imp	improve	*justice
army	invite	*stem	aim
*high	raisin	raw	revenge
people	minnow	defy	active
polish	promise	brook	purple
*working	ungracious	vile	decoy
almost	good	storm	noise
idea	*soldier	refined	table
cask	mask	thankful	unsafe
alcohol	money	fast	fame
minute	doll	tremble	strength
stain	rotten	center	scoff
nurse	*cabbage	*lamp	humble
*sour	cost	saddle	*boy
chapel	irksome	pin	interest
plant	apricot	denounce	old
sky	stone	cook	wealthy
forest	*hard	fraud	modest
companion	brute	bring	*light
repeat	escape	paint	fact
against	gain	nut	violet
finish	trifle	immense	appetite
sermon	*eagle	*dream	attraction
*earth	admit	condition	*health
conceit	snow	descend	across
crack	about	splash	piano
drag	equal	abroad	least
condemn	brown	*yellow	salmon
iron	late	deceive	price
emperor	adore	bride	garden
plead	perish	worship	scar
apart	tomb	infamous	burn
ivy	harsh	drop	ashamed
*trouble	*stomach	*bread	*bible
mouse	wasp	backwards	deny
event	unripe	pattern	quantity
claw	friend	cliff	idle
ingenious	taste	level	wash
minister	joke	body	reproach
imposo	propose	elevate	energy

XIII	XIV	XV	XVI
alarm	astonish	contented	find
distrust	vest	praise	broom
dimple	whale	pump	dagger
bounce	outfit	poem	love
twig	recover	tennis	try
indulge	embrace	guard	lazy
run	devil	cake	arrow
agree	game	calm	come
secret	towel	remorse	success
ache	faithful	play	coy
advantage	dwarf	mouth	pure
napkin	use (v)	linen	shower
hill	north	*stove	jump
shelter	thread	belt	*bitter
name	rejoice	amuse	uproar
injure	*blue	sign	catch
*memory	disaster	bag	cóntest
finger	keep	concert	empty
emblem	handsome	*long	unhappy
spool	rescue	absent	divine
unfit	audacious	maiden	*hammer
middle	cage	twist	feather
*sheep	honey	false	disorder
outrage	guess	plunge	naughty
accept	disgrace	murder	exacting
low	shark	*religion	abandon
ardent	flannel	magic	*thirsty
*bath	busy	believe	pay
emerald	unmarried	author	increase
wagon	angel	oil	*city
stun	hospital	choke	chase
gentle	secure	*whiskey	unemployed
dodge	*hungry	silver	rhyme
*cottage	jewel	noble	map
shock	nice	breast	wretched
cóntrary	carve	person	distance
hunt	provoke	influence	playful
sin	alive	magnet	impulse
asleep	*priest	glad	land
exquisite	orange	ink	moderate
sweat	battle	introduce	velvet
change	tube	profane	mix
*swift	*ocean	winter	parent
éxpert	apology	help	*square
quality	inch	repress	séparate
instant	pretty	*child	sonnet
progress	brick	field	trade
melt	verse	egg	nest
applause	*head	rat	fancy
cream	bad	mock	bench

XVII	XVIII	XIX	XX
grief	decorate	contrast	demon
parsnip	chance	unhurt	credit
purse	sack	fix	frolic
unlikely	scold	interval	include
walk	portly	fond	rascal
hod	sorrow	grain	pardon
comrade	month	mistake	soap
thought	painful	front	fear
lemon	quarrel	quart	arise
refúse	flower	lecture	cane
paper	suffer	*moon	destroy
cause	fault	usher	chart
pocket	*joy	brier	refresh
task	cab	fountain	*street
kit	discord	church	wrong
*butter	sponge	attempt	rattle
pie	mother	feed	*king
strong	*bed	tame	send
*doctor	den	medicine	glow
pig	support	glove	raft
punish	conscience	dispute	mercy
regiment	devotion	*scissors	dinner
walnut	difficult	evil	*cheese
weather	adorn	irritate	bless
*loud	immoral	advice	drive
remove	spite	neighbor	scorn
fling	brave	cravat	enter
compare	circle	*quiet	scorch
queen	lettuce	entire	expression
same	ivory	contempt	*blossom
war	urge	talk	ghost
play	imagine	touch	aboard
exercise	infinite	flag	parcel
grind	observe	anxious	dreadful
*thief	assist	hurt	small
pretend	beast	*green	oppose
knock	wheat	crumb	queer
orchard	*heavy	éscort	reduce
president	repose	notch	reward
decent	terror	bird	outlaw
croak	under	husband	*afraid
plaster	caress	control	drum
lump	lard	stable	clover
question	learn	*salt	intellect
lend	*tobacco	ornament	elbow
around	destiny	errand	milk
merry	fire	hoop	smell
*lion	consent	blush	scandal
awake	*baby	dull	happy
sacred	excellent	many	mill